THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

Section B

VOL. 65, PART 10 1 October 1952 No. 394 B

OPTICS OF METALS

The following four contributions were read at the Society's meeting at Southampton on 18th December 1951.

A New Photographic Method of Measuring the Dispersion of the Optical Constants of Metals

BY J. BOR
West Ham College of Technology

MS. received 2*nd April* 1952; *read before the Society at Southampton on* 18*th December* 1951

ABSTRACT. The information from which the optical constants are determined is produced on a photographic plate as a vertical band of variable intensity for each wavelength. This is compared microphotometrically with an adjacent band of graded intensity produced by the same wavelength.

THE basic principle of this method, as in most polarimetric methods, is the measurement of the ellipse of vibration produced when plane polarized light at an azimuth of 45° is reflected obliquely from a plane metallic surface. This requires the simultaneous determination of two independent quantities, and for this reason it has hitherto been possible to measure only one wavelength at a time. The most usual quantities are the ratio of the vertical and horizontal components ψ (fig. 1), together with the phase angle Δ between

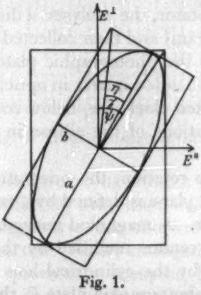


Fig. 1.

them or, as used here, the ratio of the major and minor axes of the ellipse γ and the orientation of the major axis to the vertical η. The method of calculation used is first to transform from η and γ to ψ and Δ by the geometrical transformations $\cos 2\psi = \cos 2\eta \cos 2\gamma$, $\tan \Delta = -\tan 2\eta \operatorname{cosec} 2\gamma$, and then to calculate the optical constants from the equations

$$\epsilon = n^2 - k^2 = \tan^2\phi - \frac{2\sin^2\phi\tan^2\phi\,(\cos\Delta + \sin 2\psi)}{\sin 2\psi(\cos\Delta + \operatorname{cosec} 2\psi)^2},$$

$$\sigma T = nk = \frac{\sin^2\phi\tan^2\phi\sin\Delta}{\tan 2\psi(\cos\Delta + \operatorname{cosec} 2\psi)^2},$$

where ϕ is the angle of incidence, n the refractive index, k the extinction coefficient, ϵ the dielectric constant, σ the electrical conductivity at frequency ν, τ the period of the light, i.e. $\tau = 1/\nu$.

The dispersion curve of the optical properties of metals is generally accepted as being of much greater importance than very exact measurements of one wavelength only. To make measurements over a wide spectral range it is necessary to observe one wavelength at a time over a lengthy period during which the optical properties, particularly of films, may change considerably. The apparatus to be described (Bor and Chapman 1949) enables all the information necessary for the calculation of the optical constants between approximately $0{\cdot}4\,\mu$ to $0{\cdot}65\,\mu$ to be recorded on a single photographic plate in the form of a pattern of variable density.

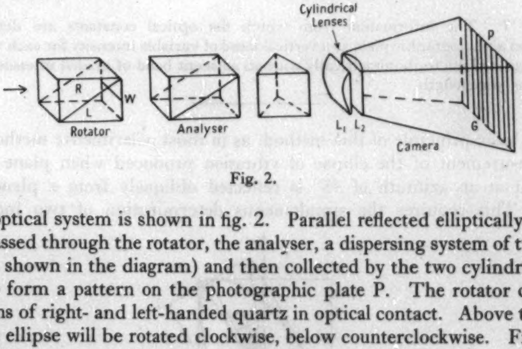
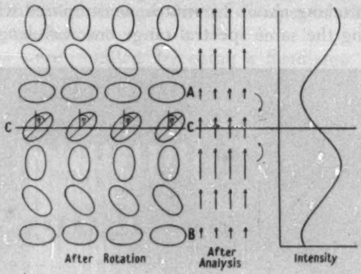


Fig. 2.

The optical system is shown in fig. 2. Parallel reflected elliptically polarized light is passed through the rotator, the analyser, a dispersing system of two prisms (one only shown in the diagram) and then collected by the two cylindrical lenses L_1, L_2 to form a pattern on the photographic plate P. The rotator consists of two prisms of right- and left-handed quartz in optical contact. Above the central plane the ellipse will be rotated clockwise, below counterclockwise. Figure 3 (*a*) indicates the relative orientations of the ellipse in the light emerging from the rotator.

If CC is the plane of no rotation, the orientation of the ellipse is given by $\eta = (\mathrm{AC}/\mathrm{AB})\pi$. This central plane is defined by a very fine wire W (fig. 2) placed immediately after the rotator. A magnified image of the wire and of the pattern of light emerging from the rotator modified by the analyser with its vibration direction vertical is thrown by the cylindrical lens L_1, of focal length 20 cm at its conjugate focus, on to a photographic plate P, the refraction taking place in a vertical direction only. Refraction in the horizontal direction only is produced by the lens L_2 of focal length 65 cm (the length of the camera), and this converges the parallel light in the horizontal plane on to the photographic plate P. This combination of lenses produces a narrow vertical band of intensity which is not uniform (fig. 3 (*b*)), but which oscillates in the ratio a^2/b^2. Hence the square root of the ratio of minimum to maximum intensity gives the ratio γ of the axes of the ellipse. Dispersion of the white light by a train of two prisms produces a spectrum crossed by a system of bright and dark bands together with the fiducial line. Adjustment is provided so that the photographic plate can be suitably tilted to allow for variations in focal length of the different wavelengths. The cylindrical lenses can be moved horizontally and can be independently rotated by slow motion about a horizontal axis. Accurate adjustment is obtained by using the line spectrum from a mercury arc, lens L_2 being adjusted before L_1 is placed in position. The adjustment about the axis is very sharp, a very small rotation of either lens producing an indistinct image.



(*a*) (*b*)

Fig. 3.

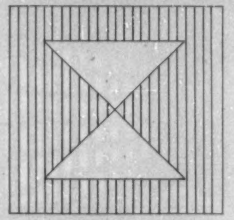


Fig. 4. Double triangular slit.

In order to translate relative blackenings on the photographic plate into true relative light intensities the following procedure was adopted. The rotator and analyser, which are mounted together as one unit, are removed and replaced by an accurate double triangular aperture (fig. 4) arranged vertically, so that its plane passes through the position previously occupied by the fiducial line. This produces a spectrum in which, owing to the action of the cylindrical lenses, the light intensity for each wavelength increases linearly with distance measured vertically from the centre. The two spectra are interleaved by placing a grid G (fig. 2), of alternate equal opaque and transparent strips, in front of the photographic plate. An exposure of about 5 seconds is first made with the rotator unit in position. The latter is then removed and replaced by the double triangular slit, the photographic plate is moved horizontally the appropriate distance in its own plane, so as to uncover the unexposed segments of the plate, and another exposure of about 5 seconds is made. Thus adjacent strips of the two spectra are formed by the same wavelength. Figure 5 shows the pattern obtained with aluminium, the whole operation taking about 2 minutes, as compared with the several hours necessary in covering the same spectral range one wavelength at a time.

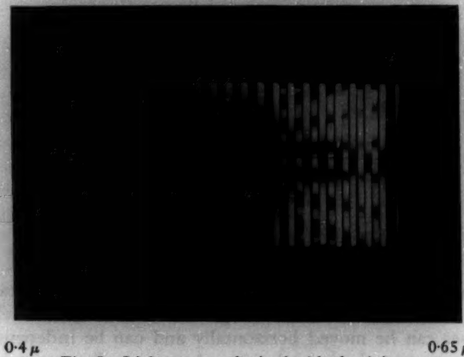


Fig. 5. Light pattern obtained with aluminium.

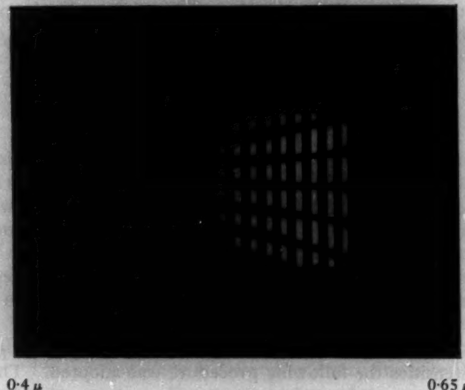


Fig. 7. Calibration light pattern.
Light and dark are reversed on the photographic plates.

A self-luminous point source of light collimated by a lens or mirror would be the ideal illuminant, since it would produce a complete pattern, and would be of uniform intensity over the field of the rotator. Since in practice a light source of finite size must be used, its permissible dimensions are governed by the following considerations: (*a*) The width must be such that the wavelength confusion in the spectrum due to horizontal divergence of the beam must be sufficiently small compared with the variation of optical constants with wavelength. (*b*) The permissible length is governed by the maximum vertical divergence of the beam. This produces confusion of the fringes because of the double refraction of a light ray which passes through quartz at a small angle to the optic axis. Calculation showed that a divergence of about $\pm \frac{1}{4}°$ would produce a negligible confusion.

These conditions were fulfilled by using a headlamp bulb with a small, straight, compact filament of size about 4×1 mm at the focus of a large aperture 50 cm achromatic lens. The wavelength confusion was no more than 25 Å in the blue and 100 Å in the red, and the uniformity as tested by a photocell along a spectrum line (rotator unit and aperture removed) was quite constant and varied by only $\frac{1}{2}\%$ at the extreme ends, which were avoided when measurements were made.

A Hilger non-recording microphotometer fitted with a 25 cm horizontal scale reading by vernier to 0·01 mm was used for measuring the intensity of the light transmitted through the plate and for fixing the positions of the maxima and minima.

Measurement of ratio of axes. Transmitted light intensities at the positions of maximum and minimum are matched with the adjacent graded intensity. Since maximum incident light on the photographic plate produces maximum blackening and minimum light transmission it follows that the ratio of the major to the minor axis is given by $(L_2/L_1)^{1/2}$ (fig. 6). Mean values are taken for all the maxima and minima for each wavelength. It is preferable to keep within the linear portion of the characteristic curve of the plates, which were Ilford special rapid panchromatic developed to a γ value of about 1. The linear region extends over a density range of 0·4 to 2·0, covering an intensity ratio of about 60 : 1, equivalent to an axial ratio of 7·7 to 1. For the metals measured the axial ratio did not exceed 3 : 1.

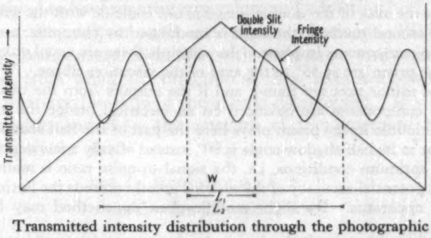


Fig. 6. Transmitted intensity distribution through the photographic plate.

Measurement of orientation. The position of the fiducial line relative to the central plane was measured on the calibration plate (fig. 7) by allowing plane polarized light vibrating in a horizontal plane to pass through the rotator and analyser, adjusting the wire to be parallel to the central minimum of transmission which coincides with the position of no rotation, and then measuring its distance from this line. The distance between the fiducial line, marked by absence of blackening, and the position of the various maxima and minima on the specimen plate (fig. 5) are measured, from which mean values of AC and AB (fig. 3 (*a*)) are obtained. In practice, readings of equal deflection are taken on both sides of the various maxima and minima.

The apparatus was tested by measuring a specimen with the author's photoelectric apparatus (Bor 1937, Bor, Hobson and Wood 1939) and then with the new apparatus. The results agreed to within the limit of experimental error throughout the spectrum range.

Measurements have been made *in vacuo* on gold, copper and aluminium films prepared by evaporation, and the results will be reported in a subsequent communication.

REFERENCES

Bor, J., 1937, *Nature, Lond.*, **139**, 716.
Bor, J., and Chapman, B. G., 1949, *Nature, Lond.*, **163**, 183.
Bor, J., Hobson, A., and Wood, C., 1939, *Proc. Phys. Soc.*, **51**, 932.

Photoelectric Analysis of Elliptically Polarized Light

By J. F. ARCHARD, P. L. CLEGG and A. M. TAYLOR
University of Southampton

MS. received 21*st March* 1952; *read before the Society at Southampton on* 18*th December* 1951

ABSTRACT. Two photoelectric methods of analysis of elliptically polarized light have been devised. One is a development of that of Kent and Lawson in which a rotating analyser yields a zero signal only when the light is circularly polarized. We use a double image prism after a rotating analyser and the two resulting modulated beams are in antiphase only when the axes of the double image prism coincide with the axes of the vibration ellipse. In our second method, the light is modulated by chopping, and a double image prism selects two components in phase. In principle these are equal only when the axes of the double image prism are at 45° to the axes of the vibration ellipse. Thus the directions of the major and minor axes are found, and if the outputs from the two photocells which receive the two components are balanced on an electrical bridge the axial ratio may be obtained. The double image prism plays here the part of the half shadow device in visual observations, but as its half shadow angle is 90° instead of only a few degrees, the photocells are used under optimum conditions, i.e. the signal-to-noise ratio is maintained as large as possible and the potential accuracy of the method greatly exceeds the limitations imposed by the rest of the apparatus. By slight modifications the method may be made partially self-recording.

Some examples are given of measurements made on metallic films undergoing surface changes.

§1. INTRODUCTION

Transmission through birefringent material, or reflection at an interface causes linearly polarized light to become elliptically polarized. Measurement of the effect is not identical in the two cases. When birefringence is not accompanied by perceptible dichroism, the phase difference Δ between the two components of light parallel and perpendicular to the two relevant axes of birefringence is the only subject of measurement. If dichroism is present, these two components suffer differential absorption, and it is necessary also to find the ratio of these components before and after transmission by the material. If, relative to the axes of birefringence, χ be the azimuth of the incident plane polarized light, then

$$T \tan \chi = \tan \chi' \qquad \ldots\ldots(1)$$

where T is the ratio of the transmission coefficients, and $\tan \chi'$ is the measured ratio after transmission.

In the case of reflection experiments, the reference planes are normal to the incident or reflected wave front, and parallel and perpendicular to the plane of incidence. The problem is then similar to the simultaneous determination of birefringence and dichroism, the two quantities of interest now being Δ, the differential phase change between the two components introduced on reflection, and ρ the ratio of their amplitude reflection coefficients, where

$$\rho = \frac{\tan \chi'}{\tan \chi} = \tan \psi. \qquad \ldots\ldots (2)$$

ψ is the so-called azimuth of restored polarization.

In what follows, we shall be concerned primarily with the reflection case, but the methods described could equally well be applied to the transmission problem.

1.1. *Visual Methods*

When birefringence alone is in question, measurement of the phase difference Δ is usually made by means of an optical compensator. The principle of some, such as that of the Babinet or Soleil type, is that the compensator introduces a phase change equal and opposite to that introduced by the specimen; the elliptically polarized light is thus reduced to plane polarized light with the same azimuth as the original incident light.

A compensator much used is that known as the Senarmont, which consists simply of a quarter-wave ($\lambda/4$) plate. One method of using it, which has similarities to the photoelectric methods discussed later, consists in allowing χ and consequently χ' to have an arbitrary value, and adjusting both the Senarmont compensator and the analyser until extinction is obtained. When this is done the axes of the compensator coincide with the major and minor axes of the vibration ellipse, from which the direction of these axes is known, and from the position of the analyser the ratio of the axes is determined. If γ be the azimuth of the compensator, and $\beta+\gamma$ the azimuth of the emergent plane polarized light, it may be shown that

$$\left.\begin{aligned} \cos 2\chi' &= \cos 2\gamma \cos 2\beta, \\ \tan \Delta &= \tan 2\beta \operatorname{cosec} 2\gamma. \end{aligned}\right\} \qquad \ldots\ldots (3)$$

Thus Δ the phase change and χ' can be calculated. The ratios of the transmission or reflection coefficients are determined by eqns. (1) or (2), from the calculated value of χ', and from the azimuth of the incident plane polarized light χ.

In the above arrangement there are three variables χ, γ and β. It is possible, and sometimes convenient, to fix one of these three. The procedure sometimes adopted is to set γ at $\pm 45°$, i.e. in the case of reflection the azimuth of the incident plane polarized light must be set so that the axes of the reflected elliptically polarized light are at 45° to the plane of incidence. This arrangement makes it possible to determine ψ and Δ directly, since in this case

$$\Delta = 2\beta, \quad \psi = \tfrac{1}{2}\pi - \chi. \qquad \ldots \quad (4)$$

Physically, these equations mean that the polarizer is set so as to compensate for the difference in the reflection coefficients, the two components (parallel and normal to the plane of incidence) being equal after reflection. The axes of the reflected ellipse are then at 45° to the plane of incidence, and their amplitude ratio is tan $\frac{1}{2}\Delta$.

1.2. *Photoelectric Methods*

Analysis of elliptically polarized light has been made in non-visible regions of the spectrum by ingenious photographic methods such as those due to Voigt (1901) and Minor (1903), to Szivessy, Dierkesmann and Munster (1933) and to Bor and Chapman (1949), but in no case is the method very rapid, for the exposure may take up to very many minutes, and much work has to be done in measuring the parameters of the patterns obtained. Sometimes this is not an insuperable objection, but if it is wished to follow the change of ellipticity as the surface undergoes physical or chemical change, then a faster method must be devised. The use of photoelectric or electronic devices seems obvious, but it appeared advantageous not to adapt the conventional visual methods to suit photoelectric detection, but to devise methods new in themselves.

Any photoelectric method should use to the best advantage the properties of the photoelectric cell. At low intensities of illumination it has little advantage over the eye, but for higher intensities it is greatly superior. A number of workers, Bruhat and Guinier (1933), Rank, Light and Yoder (1950), and Levy, Schwed and Fergus (1950), have applied photoelectric cells to polarimetry, and Bruhat and Grivet (1934) to the analysis of elliptically polarized light. Broadly speaking, these attempts, though ingenious, have not made the best use of the properties of the photocell as they have generally merely replaced the eye by a photocell in existing optical apparatus.

A promising method is that described by Kent and Lawson (1937), who used the fact that rotation of an analyser produces no change in intensity when the light is circularly polarized. This method appears very suited to the characteristics of a photocell, e.g. 'dark current' is of no importance, since the rotating analyser modulates the light beam, and the photocell is used as an a.c. device. Moreover, a narrow band amplifier used between the receiver and detector gives a greatly improved signal-to-noise ratio. We have examined this method in more detail and, as discussed below, have found a number of disadvantages which are not immediately obvious from the account given by Kent and Lawson. Moreover, the method as described by these authors must employ some form of optical compensator if it is desired to detect elliptically polarized light; the elliptically polarized light must be transformed by the compensator, not into plane polarized light as in visual methods, but into circularly polarized light.

We have, therefore, considered the possibility of analysing elliptically polarized light without the use of an optical compensator, the orientation γ and ratio tan β of the axes of the ellipse (see eqns. (3) above) being determined directly.

Two such methods, previously described in a preliminary report (Archard, Clegg and Taylor 1950), are discussed in more detail below. The essential technique involved is that a double image prism is used to select from the elliptically polarized light two components polarized in mutually perpendicular directions; each of these components is received by a separate photocell and, as the light is modulated, either mechanically or optically, the resultant signals may be matched with one another. Depending on the method of modulation, these signals may be in phase (Method A), or in anti-phase (Method B), but in both cases the matching circuit is essentially the same.

The type of matching circuit used is shown in fig. 1. The two signals are fed through cathode followers V_1 and V_3, acting partly as buffer stages, to opposite halves of the twin triode matching valve V_2. From this stage the output, zero when matched, is fed to a cathode-ray oscillograph via a tuned amplifier (Sturtevant 1947). Helical potentiometers R_1 and R_2, giving 3 600 degrees of spindle rotation, in the cathode circuits of the cathode followers, enable the signals to be matched and their ratio determined. The output of the matching circuit is fed to the tuned amplifier from either A or B, depending on whether it is desired to balance in- or out-of-phase signals.

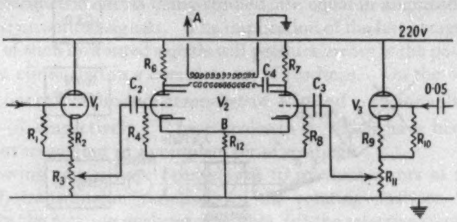


Fig. 1.

1.3. *General Discussion of Sensitivity*

The sensitivity of any photoelectric method will depend primarily on the signal-to-noise ratio. Noise arises chiefly as diode noise in the photocells, and Johnson noise in the grid resistors of the first valves. We shall make the following general assumptions: intensity of elliptically polarized light $I_0 = 10^{-4}$ lumen (mercury green light was used in our experiments); photocell sensitivity $= 30$ μA/lumen; grid resistors $R = 50$ megohms; amplifier bandwidth $= 10$ c/s.

On this basis, the noise at the grids of the first valves consists of about 4 μv of diode noise, and 3 μv of Johnson noise, or a total of about 5 μv. If it is assumed that the minimum signal-to-noise ratio should be about 4, then it follows that the minimum detectable signal is equal to that produced by a variation of light intensity $\delta I = 1{\cdot}3 \times 10^{-8}$ lumen, i.e. $\delta I/I_0 = 1{\cdot}3 \times 10^{-4}$.

This criterion will be used to derive the sensitivity of the methods discussed below. It may not, in fact, be universally applicable, but it will suffice to show that under most conditions these methods are extremely sensitive, and compare very favourably with the usual visual methods.

§ 2. THE ROTATING ANALYSER

We have applied the rotating analyser method of Kent and Lawson to the detection of circularly polarized light reflected from the metallic surface, the incident light being plane polarized. The arrangement used is shown schematically in fig. 2. The analyser was rotated at about 43 rev/sec, and the amplifier tuned to 86 c/s. The angle of incidence θ and the polarizer azimuth were adjusted until no detectable output was observed on the cathode-ray oscillograph. This occurs at the principal angle of incidence $\bar{\theta}$ (the angle of incidence at which the differential phase change Δ is $\pi/2$). There are four positions of the polarizer for zero output, two pairs each separated by an angle $2\bar{\psi}$, where $\bar{\psi}$ is the principal azimuth (the special value of ψ when $\theta = \bar{\theta}$). Thus, applied to a solid metal surface, the method enables $\bar{\theta}$ and $\bar{\psi}$ to be determined directly, and from these angles the optical constants of the metal can be determined (O'Bryan 1936).

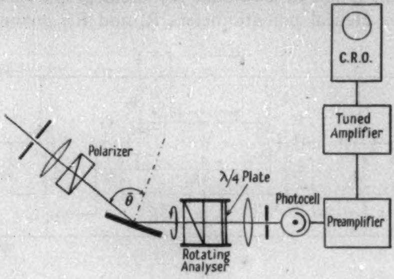


Fig. 2.

On the signal-to-noise considerations given above, the theoretical sensitivity of the adjustments corresponds to errors in ψ and Δ of about ten seconds of arc. In fact, the instrumental accuracy of the apparatus was only of the order of one or two minutes and, as discussed below, other considerations affect the final accuracy. However, the fact that internally consistent results can be obtained may be judged from four sets of consecutive readings made on an aluminium surface (table 1).

Table 1

Run no.	1	2	3	4
$\bar{\theta}$	75° 28′	75° 30′	75° 29′	75° 29′
$\bar{\psi}$	40° 50′	40° 54′	40° 55′	40° 52′

2.1. *Sources of Errors*

Consider the effect of small misadjustments from their null positions of the polarizer azimuth or of the angle of incidence (this latter is effectively a change in Δ) for the arrangement shown in fig. 1. When both adjustments are in their null positions circularly polarized light is obtained, and ideally there is no output signal. If one of the two adjustments is moved slightly, an ellipse of small ellipticity results, and a small output signal is obtained. The orientation of the ellipse and the phase of the signal differ according as to which of the two adjustments is altered. The effects may be summarized as follows :

Misadjustment	Orientation of ellipse	Output Signal
Polarizer	45° to plane of incidence	$E = \pm E_1 \cos 2\omega t$
Incidence	parallel or normal to plane of incidence	$E = \pm E_2 \sin 2\omega t$

where $2\pi\omega$ is the frequency of rotation of the analyser.

The sign of the output signal is dependent on the sign of the error of adjustment. Thus polarizer settings of $+(\bar{\chi}+\delta\bar{\chi})$ and $-(\bar{\chi}+\delta\bar{\chi})$ produce identical signals. Alternatively, for a correct setting of the polarizer azimuth but a small misadjustment of the incidence, the phase of the signal depends on which of the four positions of the polarizer azimuth is adopted. Thus a given misadjustment of incidence produces a signal $+E \sin 2\omega t$ at one polarizer position, and $-E \sin 2\omega t$ at the other.

These considerations have an important bearing on the detection of errors. A number of causes may produce variations in the light flux, or signals from the photocells, which have a component at the required frequency (86 c/s in our case). The adjustments to the apparatus now have to be made so that the required signals, due to the polarimetric effects being studied, are equal in amplitude and opposite in phase to the unwanted signals. The implication of the last paragraph is that the $\cos 2\omega t$ term of such unwanted signals will produce errors in the polarizer readings which are not eliminated in a complete set of readings. On the other hand, the $\sin 2\omega t$ component will produce errors in $\bar{\theta}$ of $+\delta\bar{\theta}$ and $-\delta\bar{\theta}$, for polarizer azimuths of $+\bar{\chi}$ and $-\bar{\chi}$ respectively. These errors in $\bar{\theta}$, which have been observed in practice, are averaged out in a complete set of readings.

The following causes have been shown to produce errors of this type : (*a*) vibration, (*b*) transmission variations in the rotating analyser, (*c*) stray light reflected from the rotating analyser assembly, (*d*) the selective photoelectric effect at the photocell, (*e*) angular deviations of the light beam produced by the rotating analyser.

It has proved extremely difficult to obtain calcite polarizing prisms which do not produce a detectable deviation of the light beam. Apart from the special consideration of unwanted signals, this deviation automatically reduces the accuracy of the measurements, because the diameter of the exit pinhole must be increased.

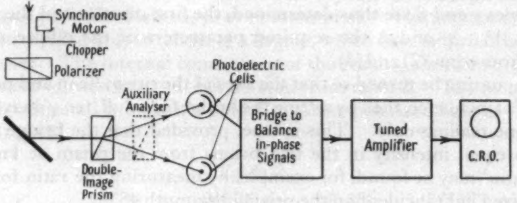


Fig. 3.

The selective photoelectric effect, combined with the semi-cylindrical shape of the photocell cathode, means that the sensitivity of the photocell is dependent on the azimuth of incident plane polarized light. In our case, the effect was such that quite small movements of the photocell made it necessary to change the polarizer azimuth by as much as one degree. The effect was eliminated by fixing a $\lambda/4$ plate in the same mounting as the rotating analyser, with its axes set at 45° to the plane of polarization of the analyser. This converted the plane polarized light emergent from the analyser into circularly polarized light.

§ 3. DIRECT ANALYSIS OF ELLIPTICALLY POLARIZED LIGHT: METHOD A

The method, an experimental arrangement of which is shown schematically in fig. 3, uses a double image prism and is similar to a visual method previously described by Ingersoll and Littleton (1910) and a photographic method due to

Pfestorf (1926). The double image prism selects two mutually perpendicular components of the elliptically polarized light, and the principle involved may be simplified as follows : (*a*) When the two components are equal, the double image prism is set at 45° to the axes of the ellipse : this determines the orientation of the ellipse. (*b*) The double image prism is then set along the axes as determined in (*a*), and the ratio of the components gives the intensity ratio of the axes of the ellipse. The ellipse is thus completely determined, except for its sense of rotation.

The method of using the apparatus will now be described in more detail. In general, no matter what the setting of the scaling potentiometers, there will be four positions of null output as the mounting is turned through 360°. These positions will be in two complementary pairs lying symmetrically with respect to the axes of the vibration ellipse. By suitably adjusting the bridge potentiometers, these null positions can be obtained at about 45° to the ellipse axes, and this can be shown to be the most sensitive position. By taking the mean value of these four positions, the direction of the axes of the vibration ellipse may be determined.

Once the axes have been located, the intensity ratio along them may be found by turning the mounting until the axes of the prism lie along those of the ellipse. If the measured ratio is p_1 in this position, then when the mounting is turned through 90°, the new ratio will be p_2, the change being due to the differences between the two circuits (amplifications, photocell sensitivities, etc.). The true ratio is given by $(p_1p_2)^{1/2}$, and the amplitude ratio, the desired quantity, by $(p_1p_2)^{1/4}$. In practice the ratios p_1 and p_2 are each found by taking the mean values of the ratios measured at positions 180° apart.

The angles γ and β are thus determined, the first directly and the second from $\tan\beta=(p_1p_2)^{1/4}$. χ' and Δ the required parameters of the ellipse may then be calculated from eqns. (2) and (3).

If the mounting be turned so that the axes of the prism lie in and perpendicular to the plane of incidence, then by setting the bridge for a null, $\tan\chi'$ may be measured from this one reading only. This is true, provided that the bridge ratio corresponding to equal intensity in the two beams from the prism be known. This 'equality ratio' may be found, for example, by measuring the ratio for a null with plane polarized light incident on the prism in azimuth 45°.

By analogy with the use of the Senarmont compensator Δ and ψ may be measured directly if the polarizer azimuth χ is adjusted till $\gamma=45°$, their values being calculated from eqn. (4).

Provided that a Wollaston rather than a Rochon double image prism be used (Archard 1950) it would be possible to match the signals by means of an auxiliary analyser, as is indicated in fig. 2, instead of by adjusting the bridge as has been described. This analyser, an ordinary polarizing prism, should be placed immediately after the double image prism, preferably in the same mounting so that its orientation with respect to the double image prism can be read directly. When the double image prism is aligned along the ellipse axes, the axial ratio can be measured from that orientation of the analyser which results in a null output.

If the electrical circuits associated with the two photocells were adjusted for approximate equality, the average value obtained from the four positions of the auxiliary analyser would suffice to give the axial ratio. This analyser could be set at 45° in the determination of the orientation γ by the procedure described above.

3.1. *Sensitivity*

Assume that the method described above has been set up, using an auxiliary analyser. It is required to measure the orientation γ and ratio $\tan\beta$ of the axes of the ellipse, these being given directly from the settings of the double image prism and auxiliary analyser respectively. The signals resulting from angular errors of adjustment $\delta\gamma$ and $\delta\beta$ may be regarded as being due to light intensities δI_γ and δI_β respectively. Then, it may be shown that

$$\frac{\delta I_\gamma}{I_0} = \cos 2\beta\delta\gamma, \qquad \frac{\delta I_\beta}{I_0} = \sin 2\beta\delta\beta \qquad \ldots\ldots(5)$$

where I_0 is the intensity of the elliptically polarized light being analysed.

Once again, applying the criterion of § 1.3, it may be seen that under a wide range of conditions the theoretical sensitivity in the measurement of γ and β is much less than one minute of arc.

Usually the quantities required are ψ and Δ, and therefore it is necessary to convert these errors in γ and β to errors in ψ and Δ, but, except under the most unfavourable conditions, they will be of the same order. In the special case where γ is set at 45°, it has been shown that $\Delta = 2\beta$, and ψ is given directly from the polarizer setting. In this case, the errors in ψ and Δ are given by the equations

$$\frac{\delta I_\gamma}{I_0} = 2\operatorname{cosec} 2\psi\delta\psi, \qquad \frac{\delta I_\beta}{I_0} = \tfrac{1}{2}\sin\Delta\delta\Delta. \qquad \ldots\ldots(6)$$

3.2. *Results*

Though, as has been shown, the method is very sensitive, a high experimental accuracy has not yet been attained, or even attempted. The angular scales in our apparatus were divided only to tenths of a degree, which gave sufficient accuracy for present purposes. The internal consistency of the results, a measure of the sensitivity of the method, was good, as can be seen from table 2. The results refer to values for light reflected from a thin gold film, plane polarized light at 45° azimuth being incident at 45°.

Table 2

γ	24° 40′	24° 37′	24° 38′	24° 39′
β	12° 05′	12° 05′	12° 05′	12° 06′
ψ	26° 46′	26° 43′	26° 44′	26° 45′
Δ	30° 36′	30° 39′	30° 38′	30° 37′

Much work has been done with the apparatus in determining the values of ψ and Δ as a function of thickness for thin films of various metals evaporated and measured *in vacuo* (Clegg 1952). As an example of measurements within the scope of the apparatus, fig. 4 shows the variation of ψ and Δ with thickness for a copper film; the thickness was determined by means of the microbalance (Clegg and Crook 1952), the density of the film being assumed to be that of bulk copper. Figure 5 shows an automatic recording of the oxidation of a thin copper film which was obtained as follows: after deposition of the copper film the bridge was balanced so as to measure $\tan\chi'$ directly as has already been indicated; air was then admitted to the system and the film partly oxidized, thus altering the value of χ'. The bridge, which was not touched, was then no longer balanced, and it was this out-of-balance signal that was suitably recorded.

§ 4. DIRECT ANALYSIS OF ELLIPTICALLY POLARIZED LIGHT: METHOD B

In Method A, described above, the bridge is balanced for signals in phase, and two separate adjustments of the apparatus are necessary to obtain the orientation γ and amplitude ratio tan β of the axes of the ellipse. The method described below utilizes the fact that the bridge must be balanced both in amplitude and phase. In this way it is possible to obtain γ and θ from a single adjustment of the apparatus, the two variables being alternately adjusted until no output is obtained.

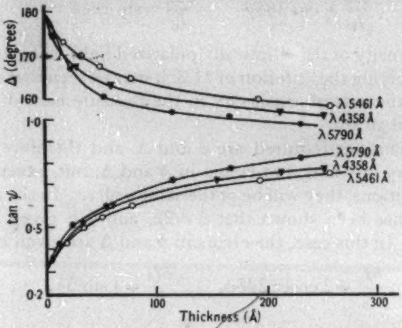


Fig. 4. Copper films of increasing thickness measured at 45° angle of incidence for various wavelengths. The evaporation rate was 1 Å per second.

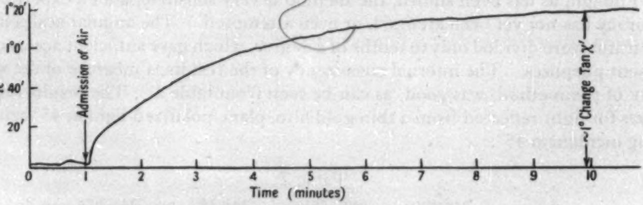


Fig. 5. Automatic recording of oxidation of an evaporated film of copper.

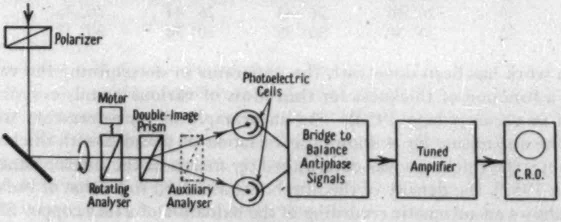

Polarizer
Motor
Double-Image Prism
Photoelectric Cells
Rotating Analyser
Auxiliary Analyser
Bridge to Balance Antiphase Signals
Tuned Amplifier
C.R.O.

Fig. 6.

The principle of the method is illustrated in the schematic diagram fig. 6. The general arrangement is similar to that employed in Method A, except that the light falling on the photocells is modulated by a rotating analyser placed immediately before the double image prism. The modulations of the two beams falling on the photocells contain components of frequencies $2f$ and $4f$, where f is the frequency of rotation of the rotating analyser. The $4f$ components are rejected by a filter, and the two photocell outputs thus effectively consist of signals of frequency $2f$. The complete analysis shows that the two signals have, in general, an arbitrary phase difference which depends on the mutual orientation of the axes of the vibration ellipse, and the axes of the double image prism. When, and only when, the axes of the ellipse coincide with the axes of the double image prism are the two signals in anti-phase. When this adjustment has been made, the ratio of the two signals is equal to the intensity ratio of the axes of the ellipse. Thus, by correctly orienting the double image prism to coincide with the axes of the ellipse, and balancing the signals by means of the bridge or auxiliary analyser, the orientation and ratio of the axes of the ellipse can be determined.

4.1. *Sensitivity and Accuracy*

For comparison with the discussion of Method A in § 3.1 above, it will again be assumed that an auxiliary analyser is used. A detailed analysis of Method B shows that its sensitivity is identical with that of Method A. Thus eqns. (5) and (6) apply also to this method. The final accuracy will, however, be affected by the additional sources of errors inherent in the use of a rotating analyser which were discussed in § 2.1. When the method is applied to the direct determination of ψ and Δ we have a situation analogous to that discussed in § 2.1. These special effects result in errors in the measured value of ψ, whereas such errors in Δ are eliminated in a complete set of readings. Similarly, when the method is used to determine γ and β the resultant errors in the measurement of γ are not eliminated

4.2. *Results*

Some results are illustrated in table 3, from measurements on an aluminium surface. The apparatus was arranged to measure ψ and Δ directly, ψ being obtained from the polarizer adjustment, and Δ from the setting of the helical potentiometers in the bridge circuit.

Table 3

θ	80° 00′	76° 28′	75° 00′	70° 00′	65° 00′	60° 00′	55° 00′	50° 00′
ψ	41° 49′	41° 20′	41° 28′	41° 44′	42° 14′	42° 43′	43° 06′	43° 33′
Δ	70° 08′	90° 00′*	96° 10′	114° 42′	128° 36′	139° 44′	148° 16′	155° 12′

* Values of $\overline{\theta}$ and $\overline{\psi}$ measured using rotating analyser as described in § 2.

§ 5. CONCLUSIONS

The work described in this paper has demonstrated the feasibility of new methods for the analysis of elliptically polarized light. One of the main advantages of these methods is that they are applicable to a wide spectral range, provided polarizing devices and radiation detectors are available. Moreover, one very important advantage which is gained from the elimination of optical compensators is that the apparatus does not require recalibration at each wavelength used.

In our work, the accuracy has been limited mainly by instrumental considerations (e.g. the divided circles available in our apparatus), and it is not yet known whether the extremely high theoretical accuracy suggested by the sensitivity calculations can be achieved in practice. On accuracy considerations alone, the discussion given above suggests that the use of a rotating analyser is undesirable, and the type of system used in Method A seems preferable on practically all counts.

It may be advantageous to list here some of the experimental details which require careful attention if the highest accuracy is to be attained.

(i) The linearity of the photocells and their associated circuits are important when the ratio is obtained electrically. On the other hand, if an auxiliary analyser is used it would appear that only careful matching of the two circuits is required.

(ii) The circuits, as a whole, must be stable over periods of at least a few minutes, during which time γ and β may be determined.

(iii) The light beam must be of uniform intensity and suffer no translational movement as the mounting is rotated; alternatively, the photocathodes must be of uniform sensitivity.

The methods described are certainly as rapid as normal visual methods of comparable accuracy, and outside the visible spectral range they are much more rapid than previous methods. Figure 5 demonstrates the way in which rapid changes may be followed by the development of automatic recording devices. Such devices may, for example, be used to record changes in Δ provided that no dichroism is present. When the axes of a birefringent material are at 45° to the incident light and if a quarter-wave plate is used to reduce the elliptical vibration to a linear one, then a small change in Δ will simply rotate this plane of polarization. This change in orientation may be readily recorded by means of the out-of-balance signal, which can be shown to be directly proportional to the change in Δ over quite a reasonable range $\delta\Delta$. The method is used to its best advantage under these conditions of maximum illumination. The application to accurate rapid polarimetry where β is zero and γ alone is to be measured is obvious, and would enable optical rotations to be determined to greater accuracy than by visual means.

ACKNOWLEDGMENTS

The authors wish to acknowledge the assistance of Professor E. E. Zepler in the discussion of the electronic problems. Thanks are due to the Royal Society and the Carnegie Trust for grants for apparatus. Part of the work described in this paper has been included in theses submitted by two of us (J.F.A. and P.L.C.) for the Ph.D. degree of the University of London. Both these authors wish to thank the Ministry of Education for grants which enabled them to carry out this work.

REFERENCES

Archard, J. F., 1950, *J. Sci. Instrum.*, **27**, 238.

Archard, J. F., Clegg, P. L., and Taylor, A. M., 1950, *Research*, **3**, 339.

Bor, J., and Chapman, B. G., 1949, *Nature, Lond.*, **163**, 182.

Bruhat, G., and Grivet, P., 1934, *C. R. Acad. Sci., Paris*, **199**, 852.

Bruhat, G., and Guinier, A., 1933, *C. R. Acad. Sci., Paris*, **196**, 762.

Clegg, P. L., 1952, *Thesis*, University of London.

Clegg, P. L., and Crook, A. W., 1952, *J. Sci. Instrum.*, **29**, 201.

Ingersoll, L. R., and Littleton, J. T., 1910, *Phys. Rev.*, **31**, 489.

Kent, C. V., and Lawson, J., 1937, *J. Opt. Soc. Amer.*, **27**, 117.

Levy, G. B., Schwed, P., and Fergus, D., 1950, *Rev. Sci. Instrum.*, **21**, 693.

Minor, R. C., 1903, *Ann. Phys., Lpz.*, **10**, 581.

O'Bryan, H. M., 1936, *J. Opt. Soc. Amer.*, **26**, 122.

Pfestorf, G., 1926, *Ann. Phys., Lpz.*, **81**, 906.

Rank, D. H., Light, J. H., and Yoder, P. R., 1950, *J. Sci. Instrum.*, **27**, 270.

Sturtevant, J. M., 1947, *Rev. Sci. Instrum.*, **18**, 124.

Szivessy, G., Dierkesmann, A., and Munster, C., 1933, *Z. Phys.*, **82**, 279.

Voigt, W., 1901, *Phys. Z.*, **2**, 303.

The Absorption of Light by Noble Metals and its Relation to the van der Waals Contribution to the Cohesive Energy

BY J. FRIEDEL
H. H. Wills Physical Laboratory, University of Bristol

MS. received 21*st March* 1952; *read before the Society at Southampton on* 18*th December* 1951

ABSTRACT. Noble metals show absorption in the blue or ultra-violet; to this is due the colour of copper and gold. As was pointed out by Mott, the main absorption seems due to internal photoelectric effect, with excitation of a d electron. This absorption could be analysed—as for the alkalis—as a transition between two bands, described by Bloch orbitals, here d and s. But we must take into account the correlation between the hole created in the nd shell and the Fermi electrons. We can assume that the hole is screened by a $(n+1)$s electron bound to it, so that absorption should begin for the frequency of formation of the excited $nd^9(n+1)$s configuration of the monovalent ion. Agreement with experiment obtained using this model is satisfactory both in the pure metals and their alloys. The same model may be used for x-rays, and optical absorption energy is related to the energy separation between the s and d parts of K and $L_{II, III}$ emission bands. An order of magnitude for the van der Waals interaction energy between d shells could perhaps be deduced from the structure on the short wavelength side of the absorption peak.

§ 1. OPTICAL ABSORPTION IN ALKALI AND NOBLE METALS

THE small optical absorption of the alkali metals is successfully explained in terms of the conduction electrons. These electrons may absorb energy in two ways (Mott and Jones 1936). First they oscillate in the alternating field of the incident photon, and these oscillations are damped by the resistivity of the metal. Also, if the energy of the photon is large enough, some conduction electrons may make a transition to a state of an upper band in the crystal. The two mechanisms seem to explain satisfactorily the optical absorption in alkali metals, as was recently shown by Butcher (1951).

The noble metals on the other hand show a strong absorption in the optical or near ultra-violet region (fig. 1, taken from Seitz 1940, after Minor 1903 and Meier 1910). This absorption is more than ten times as large as for the alkalis and we must look for another mechanism to explain it.

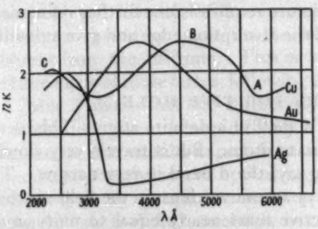


Fig. 1. Optical absorption $n\kappa$ in noble metals (*after* Seitz).

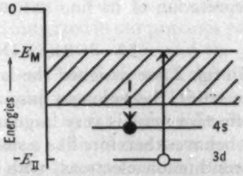


Fig. 2. Electronic transitions for the edge of the absorption peak in copper.

Mott and Jones (1936) first pointed out that the three noble metals have d shells with not too high ionization energies, and that the strong optical absorption was probably related to some excitation of the d shells.

Thus in a simple description (fig. 2), we take an electron from a d shell and put it at the top of the conduction band. We know from Fuchs' study (1935) that the d shells overlap very little in copper. The energy required to extract a d electron and send it *outside the metal* (potential 0 in fig. 2) would be much the same as if the Cu$^+$ ion was *in vacuo*, but for the fact that the ion is in the (nearly constant) field due to the lattice (other Cu$^+$ ions and Fermi gas). Thus the energy required is the difference between the second ionization potential E_{II} and a coulomb term E_c. We do not, however, send the electron outside the metal, but to the top of the conduction band. So we gain the negative energy $-E_M$ corresponding to the top of the conduction band. E_M is obtained by the approximate relation (Mott and Jones 1936): $E_M \simeq E_s + E_I - 0{\cdot}4\, E_F$ where E_s, E_I and E_F are the sublimation energy, the first ionization potential and the Fermi energy of the metal.

The total energy required is finally:

$$h\nu \simeq E_{\mathrm{II}} - E_c - E_M. \qquad \ldots\ldots(1)$$

E_c is the difference of the coulomb interactions with the rest of the lattice of a Cu$^+$ and a Cu^{2+} ion. For Cu$^+$, contributions by the surrounding atomic polyhedra cancel out very nearly; there remains only the interaction of Cu$^+$ with the part of the Fermi gas enclosed in its own atomic polyhedron. The same is true for Cu^{2+} *if we assume that the local density of the Fermi gas remains unaltered during the absorption.* Computing the fields of Cu$^+$ and Cu^{2+} ions by Fermi's approximation of the Thomas–Fermi method (Gombás 1949), we find then the values of E_c listed in table 1 for copper; those for silver and gold are also given. The computed values of the absorption edge $h\nu$ are somewhat larger than the experimental ones.

Table 1. Optical Absorption Edges $h\nu$ (in ev) computed from eqn. (1)

Metal	E_{II}	E_M	E_c	$h\nu_{\mathrm{comp}}$	$h\nu_{\mathrm{exp}}$
Cu	20·34	8·38	8·04	3·92	$2{\cdot}1_5$
Ag	21·4	8·33	7·36	5·71	4·0
Au	19·95	10·87	6·66	2·42	~2·3

This could be explained by taking into account the width of the d band.

As, however, energy E_c is large, the ionization of the absorbing ion must perturb strongly the Fermi gas, and our picture requires some further refinement. This will improve the computed values of the absorption edge and give a possible interpretation of its fine structure.

§2. SCREENING OF THE POSITIVE HOLE

In fig. 2, we assigned the hole to the d shell of a definite atom. This is not quite true: the hole may jump from atom to atom. But it travels very slowly; its effective mass is very large or, we may say, the d band is very narrow. The hole behaves therefore like a slow and heavy atomic nucleus, in the field of which the conduction electrons, with their effective mass nearly equal to unity, must at every instant be in approximate equilibrium.* They must in particular *screen* the coulomb field of the hole; for it is well known that no such field may exist in a metal in equilibrium. And the screening most likely is by a bound 4s electron (fig. 2).

---

* We apply therefore the Franck–Condon principle.

We picture therefore the edge of the absorption peak in copper as follows: a 3d electron makes a transition to the top of the conduction band, leaving a hole in the d shell; the hole is screened by a 4s bound electron provided by the conduction band.*

The 4s electron does not overlap the neighbouring ions much; so its energy is not very different from that of the 4s electron of a free Cu$^+$ ($3d^9\,4s$) ion. The energy of the conduction electrons on the other hand is probably not very much perturbed by the replacement of a normal $3d^{10}$ Cu$^+$ ion by an excited $3d^9\,4s$ one. The minimum energy required for the absorption must therefore be near to the energy of excitation $3d^{10}\rightarrow 3d^9 4s$ of a free Cu$^+$ ion.

§ 3. COMPARISON WITH EXPERIMENT

Table 2 gives the comparison for copper, silver and gold. Experimental values are taken from Seitz (1940), after Minor (1903) and Meier (1910). E_1, E_2, E_3 and E' are the $n\mathrm{d}^{10}\rightarrow n\mathrm{d}^9(n+1)\mathrm{s}$ excitation energies for a free monovalent ion, as measured in spark spectra (Cu and Ag, Bacher and Goudsmit 1932; gold, Platt and Sawyer 1941). The four values correspond to the various triplet and singlet configurations of the final state, $^3D_3$, $^3D_2$, $^3D_1$ and $^1D_1$ respectively. As we shall explain in a further paper, the excitation to the triplet state, which leads to a configuration with a non-zero total spin, is probably weaker than that to the singlet state. We should therefore expect an absorption increasing rather slowly up to a maximum corresponding to E'.

Table 2

	Experimental Values (ev)			Theoretical Values (ev)				
	Edge A	Peak B	B−A	E_1	E_2	E_3	E'	$E'-E_1$
Cu	$2{\cdot}1_5$	2·7	$0{\cdot}5_5$	2·70	2·82	2·96	3·25	0·55
Ag	4·0	4·7	0·7	4·83	5·02	5·40	5·70	0·87
Au	~2·3	$3{\cdot}2_5$	~1·0	1·85	2·18	3·42	3·66	1·81

The agreement with experiment of these roughly estimated values is surprisingly good. The energy terms neglected here are mainly an attraction of the bound state by the neighbouring d shells, and its exchange interaction with the conduction electrons, both of which tend to reduce the energy of absorption (cf. Friedel 1952). In contrast with what happens in alkali metals, both terms should be small, for the $3d^9 4s$ ion has a spatial distribution not very different from the $3d^{10}$ one. This is in agreement with a very rough estimation we have made of those terms, following the method indicated in our previous paper.

Table 2 shows that the difference in energy between the edge and the peak of absorption, given theoretically by $E'-E_1$, increases regularly with the atomic number, as one should expect from a triplet–singlet separation. Gold on the other hand absorbs in the optical region, and not in the far ultra-violet as one would expect by comparison with copper and silver; this is due to a property of the $5d^{10}$ shell of Au$^+$, probably in connection with the existence of 4f excited states of low energy.

---

* The 4s bound state may possibly be actually replaced by a heaping up of charge in the conduction band, as described by Friedel (1952). This point is not of great importance here; for we showed that the heaping up of charge has then nearly the same spatial distribution and average energy as the bound state would have if it existed.

In conclusion, the main term in the energy of the absorption edge seems to correspond to the $nd^{10} \rightarrow nd^{9}(n+1)$s excitation of the free monovalent ion. This is independent of the position of the conduction band. Thus one should expect the absorption edge to be very little displaced when alloying elements of valency other than unity are added to the noble metal. This seems in agreement with experiment (aluminium in copper, McPherson 1940, zinc in copper, Lowery, Wilkinson and Smare 1937, nickel in copper, Bor, Hobson and Wood 1939). The small displacement of the absorption edge may be due to changes in coulomb and exchange interactions of the bound state with the conduction electrons or the neighbouring atomic cores.

§4. DOUBLET STRUCTURES IN X-RAY EMISSION BANDS

The same sort of description is also applicable to any x-ray transition. At the end of a K emission process in copper for instance, the hole in the K shell has been filled by an electron coming either from the conduction band (fig. 3 (*a*)) or from the 3d shell (fig. 3 (*b*)); in that case, the hole in the 3d shell must be screened as before by a 4s bound state produced by the conduction band. Configurations (*a*) and (*b*), drawn in fig. 3 for the transitions of highest energy (emission edges), have obviously different energies; (*a*) and (*b*) correspond respectively to the beginning and the end of the optical transition described in fig. 2, and their energies differ by that necessary for the optical transition.

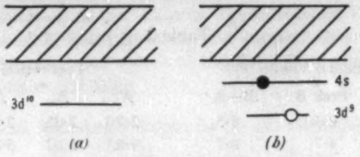


Fig. 3. End of CuK emission.

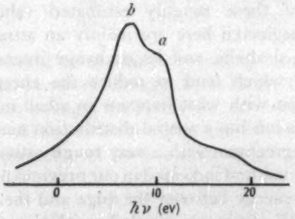


Fig. 4. CuKβ emission (*after* Bearden and Friedman).

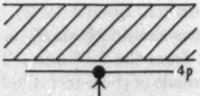


Fig. 5. Structure in Cu optical absorption.

We must therefore expect an emission band with two edges, as is indeed observed (fig. 4, from Bearden and Friedman 1940), the difference of energy between the two edges being equal to the energy of the edge of the optical absorption peak. The passage from the $(1s3d^{10})Cu^{+}$ ion to the $(1s^{2}3d^{9})$ one is made possible by the presence of the conduction electrons, in the way described above for optical transitions. Some direct quadrupole transition may however occur in the emission of type (*b*), and explain its rather symmetrical shape.

This makes difficult an accurate measurement of the separation between the edges. Bearden and Friedman's data give about 3 ev, which must be compared with the 2·1 ev obtained in optical absorption (table 1).

The Cu $L_{II, III}$ emission, where the description is analogous, gives a similar doublet, with a separation of about 2·5 to 3 ev (Cauchois 1950, 1952, cf. also Saur 1936 and Farineau 1939). Here a direct transition of the excited $(2p^5 3d^{10})Cu^+$ ion to its stable state $2p^6 3d^9$ is possible, and gives a very strong (*b*) peak.

§ 5. SCREENING BY p ELECTRONS AND VAN DER WAALS INTERACTION BETWEEN d SHELLS

We have assumed so far a screening of the *n*d hole by a $(n+1)$s bound state. But screenings by higher excited states are *a priori* possible, $(n+1)$p being the most favourable energetically.

We must therefore expect a structure on the short wavelength side of the main absorption peak. Meier's measurements on copper (1910) show a secondary peak with a maximum probably near to 2 400 Å or 5·15 ev. The energy for the $3d^{10} \rightarrow 3d^9 4p$ excitation of Cu^+ *in vacuo* is definitely larger, from 8·2 to 9·1 ev (Bacher and Goudsmit 1932). However, the 4p bound state, although it has a nodal plane, probably overlaps more than the 4s state does the neighbouring copper ions. The resulting attraction by the neighbouring ions would explain the difference from 8·2 to 5·1 ev.

If we neglect the absorption of higher frequencies the van der Waals interaction energy due to the d shells is given by (Mott and Jones 1936):

$$E_v \simeq -\frac{3{\cdot}63}{\Omega^2}\,\frac{3}{4}\left(\frac{\mathbf{e}^2\hbar^2}{\boldsymbol{m}}\right)^2\left[\frac{f^2}{(\boldsymbol{h}\nu)^3}+2\,\frac{2ff'}{\boldsymbol{h}\nu\,\boldsymbol{h}\nu'(\boldsymbol{h}\nu+\boldsymbol{h}\nu')}\right]. \quad \ldots\ldots(2)$$

Here $\boldsymbol{h}\nu$ and $\boldsymbol{h}\nu'$ are the energies of the $3d^{10} \rightarrow 3d^9 4p$ and $3d^9 4s$ transitions respectively, f and f' their oscillator strengths and Ω the atomic volume. $3d^{10} \rightarrow 3d^9 4p$ is the first direct excitation allowed for the d shell (fig. 5). So the first term in (1) gives the interaction between d shells, the second the polarization by a d shell of the conduction band on top of a neighbouring d shell. The term in $f'^2/(\boldsymbol{h}\nu')^3$ does not appear in (2), for it would correspond to the polarization of the conduction electrons by themselves, for which a formula of type (2) is certainly not valid and should be replaced by a direct study of correlation.

For copper, $\boldsymbol{h}\nu = 5{\cdot}1_5$ ev and $\boldsymbol{h}\nu' = 2{\cdot}7$ ev; hence $E_v \simeq -1{\cdot}7_3 f(f+5f')$. The exact evaluation of f and f' would be very difficult. But we should obtain a very rough order of magnitude by considering the absorption peaks as broadened lines; f is then given in terms of the maximum absorption $n\kappa$ (Heitler 1936) by:

$$f = \frac{3\boldsymbol{m}}{4\pi\hbar^2\mathbf{e}^2}\,\frac{\boldsymbol{h}\nu}{N}\,\gamma n\kappa$$

where $\boldsymbol{h}\nu$ is the energy of the absorption peak, γ its width and N the number of atoms per cm³. From fig. 1, we deduce the reasonable data: $n\kappa \simeq n\kappa' \simeq 2{\cdot}0$ and $\gamma \simeq 2{\cdot}1$; $\gamma' \simeq 1{\cdot}35$ ev. Thus $f \simeq 0{\cdot}43$; $f' \simeq 0{\cdot}27_5$ and $E_v \simeq -1{\cdot}3_5$ ev.

This is only a very rough order of magnitude. But its large value may explain the discrepancy between the cohesive energy of copper and Fuchs' computations (1935). Adding to Fuchs' result ($E_F = -1{\cdot}43$ ev/atom) a correction due to Hsu (1951) for coulomb interaction with the conduction electrons of the slightly interpenetrating d shells ($E_H = -0{\cdot}5$ ev/atom) we have $E_F + E_H = -1{\cdot}9_5$ ev/atom, which differs from the experimental value (Kubaschewski and Evans 1951) $E_c = -3{\cdot}54 \pm 0{\cdot}0_9$ ev/atom by an amount of the order of E_v: $E_c - (E_F + E_H) \simeq -1{\cdot}6$ ev/atom. A further investigation of the absorption of noble metals in the ultra-violet would therefore be of interest, to check our interpretation of the structure and eventually to obtain a more precise estimate of the van der Waals interactions.

ACKNOWLEDGMENTS

The author wishes to thank Professor N. F. Mott, who introduced him to the subject, for constant interest and advice, Dr. F. Abeles for helpful discussion, and the Direction des Mines for allowing him to stay in Bristol. Section 1 has been modified according to the comments of one referee.

REFERENCES

Bacher, R., and Goudsmit, S., 1932, *Atomic Energy States* (New York and London: McGraw-Hill).

Bearden, J. A., and Friedman, H., 1940, *Phys. Rev.*, **58**, 387.

Bor, J., Hobson, A., and Wood, C., 1939, *Proc. Phys. Soc.*, **51**, 942.

Butcher, P. N., 1951, *Proc. Phys. Soc.* A, **64**, 765.

Cauchois, Y., 1950, Rep. Madison Conf.; 1952, *Phil. Mag.* (in the press).

Farineau, J., 1939, *J. Phys. Radium*, Ser. 7, **10**, 327.

Friedel, J., 1952, *Phil. Mag.*, **43**, 153.

Fuchs, K., 1935, *Proc. Roy. Soc.* A, **151**, 585.

Gombás, P., 1949, *Die statistische Theorie des Atoms* (Wien: Springer), p. 48.

Heitler, W., 1936, *Quantum theory of Radiation* (Oxford: University Press), p. 115.

Hsu, Y. C., 1951, 21st Quart. Rep. Study of Metals, Chicago, May–July (to be published in *Phys. Rev.*).

Kubaschewski, O., and Evans, E. Ll., 1951, *Metallurgical Thermochemistry* (London: Butterworths Scientific Publications).

Lowery, H., Wilkinson, H., and Smare, D. L., 1937, *Proc. Phys. Soc.*, **49**, 345.

McPherson, L., 1940, *Proc. Phys. Soc.*, **52**, 210.

Meier, W., 1910, *Ann. Phys., Lpz.*, **31**, 1017.

Minor, R. S., 1903, *Ann. Phys., Lpz.*, **10**, 581.

Mott, N. F., and Jones, H., 1936, *Metals and Alloys* (Oxford: University Press).

Platt, J. R., and Sawyer, R. A., 1941, *Phys. Rev.*, **60**, 866.

Saur, E., 1936, *Z. Phys.*, **103**, 421.

Seitz, F., 1940, *Theory of Solids* (New York and London: McGraw-Hill).

The Optical Constants of Thin Metallic Films Deposited by Evaporation

By P. L. CLEGG
University of Southampton

MS. received 21*st March* 1952; *read before the Society at Southampton on* 18*th December* 1951

ABSTRACT. The optical properties of evaporated films of silver, gold, tin and indium have been determined by a photoelectric method. The influence of factors concerned with the evaporation procedure has been investigated in the case of the silver films. The experimental results are accounted for by the agglomerated state of the films, and the application of Maxwell Garnett's theory is shown to lead to qualitative agreement with the experimental values.

§ 1. INTRODUCTION

The measurements detailed below were made with the photoelectric apparatus that has already been described as Method A by Archard, Clegg and Taylor (1952). The metallic films were produced by downward evaporation on to a clean, wedged, glass substrate at a pressure of about 2×10^{-5} mm of mercury. The required optical properties of the film were then determined without breaking the vacuum; these were $\tan \psi = r_p/r_s$, the ratio of the Fresnel reflection coefficients of light polarized in and perpendicular to the plane of incidence; and $\Delta = \delta_p - \delta_s$, the differential phase change of these components on reflection from the film.

A schematic diagram of the apparatus used is shown in fig. 1 and is largely self-explanatory. Measurements were made at angles of incidence of 45° and 65°. Without breaking the vacuum this angle could be varied by rotating the polarizer and analyser arms, in opposite senses, about the axes DD and BB. A 240-watt box-type mercury lamp was used as a source.

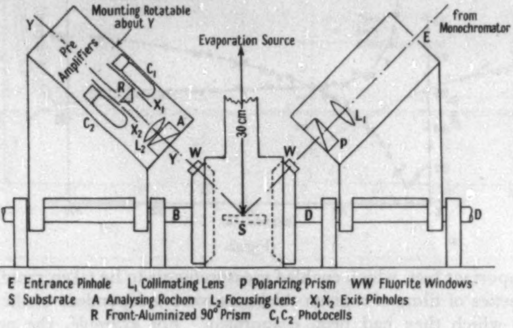

Mounting Rotatable about Y
Y
Pre Amplifiers
C_1
R
X_1
C_2
X_2
A
L_2
Y
W
Evaporation Source
30 cm
from Monochromator
E
L_1
P
W
B
B
S
D
D
E Entrance Pinhole
L_1 Collimating Lens
P Polarizing Prism
WW Fluorite Windows
S Substrate
A Analysing Rochon
L_2 Focusing Lens
$X_1 X_2$ Exit Pinholes
R Front-Aluminized 90° Prism
$C_1 C_2$ Photocells

Fig. 1.

The mass per square centimetre of the deposited film was directly measured by weighing with the vacuum microbalance, described by Clegg and Crook (1952), placed adjacent to the substrate. For convenience this measure has been transformed into a thickness, the density of the film being assumed equal to that of the bulk material; this thickness is known as the 'bulk thickness' of the film. For thicker films this value was checked, after their removal from the vacuum chamber, by a multiple-beam interferometric method.

§ 2. EXPERIMENTAL RESULTS

The first metal to be extensively studied was silver, chosen because it is inert, and is therefore not likely to undergo any chemical change in the atmosphere of the vacuum system. Various factors connected with the production of the films were altered so as to determine those that were most important. The only factors to have any appreciable effect on the resultant optical properties were the evaporation rate, the temperature of the substrate and the cleanliness of the substrate.

It was thought that the different topographies of plate glass, fire-polished glass and an optical flat might have influenced the growth of the films, but there was no difference between the optical properties of films deposited on various types of glass substrate. There was also no appreciable difference between films deposited from tantalum, molybdenum, or plantinized tungsten filaments: it was therefore decided to use the latter for the subsequent depositions.

In fig. 2 are shown some results obtained for various evaporation conditions at a medium rate of about 2 Å per second. The scatter of the experimental points for any given run is small, and the scatter between runs gives an indication of the experimental accuracy of these and the other experimental results quoted in this paper.

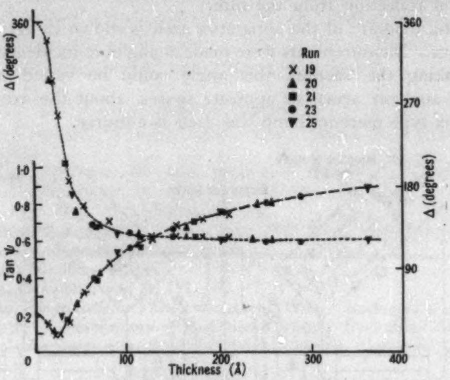
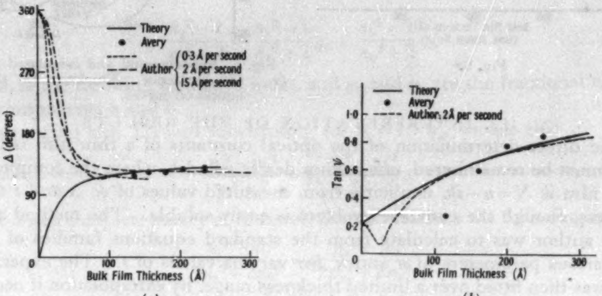


Fig. 2.

An important fact, which enabled measurements to be taken rapidly, was that the properties of films of a given thickness were independent of the number of stages in which they had been evaporated. For example, the properties of films 150 Å thick built up in 2, 6 or 8 steps were the same. Although this applied only to films deposited at, or slower than, an evaporation rate of 2 Å per second it enabled the optical properties of the films built up *in vacuo* to be determined as functions of film thickness.

The original measurements on silver films deposited at a rate of about 2 Å per second were made at an angle of incidence $\theta_1 = 65°$ for a wavelength of 5461 Å. The results obtained are compared in figs. 3 (*a*) and 3 (*b*) with those obtained by Avery (1949); the theoretical curves in these figures were calculated from the standard formulae for thin films as given by Drude (1894) and Lord Rayleigh (1902), the optical constants of the film (homogeneous) being assumed to be those of bulk silver. The disagreement between the author's results for Δ and those expected theoretically and obtained, down to about 100 Å, by Avery led to further experiments in which the evaporation rate was varied. These results are also shown in fig. 3 (*a*), and it may be seen that an increase in the evaporation rate extends the agreement between theory and experiment to smaller film thicknesses, though for the thinnest films there is still a major discrepancy.

In fig. 4 are shown for various wavelengths the detailed results for silver at $\theta_1 = 65°$, the evaporation rate being 2 Å per second; it will be noted that there is a complete change in the character of the results, so far as the values of Δ are concerned, between $\lambda = 5461$ Å and 4358 Å. A similar change is also apparent in the results shown in fig. 5 for an angle of incidence $\theta_1 = 45°$ at the same evaporation rate; the rise in the value of Δ above 180° is again quite unexpected; theory, as was the case for $\theta_1 = 65°$, predicts the very reverse. From a knowledge of the optical constants of almost any bulk metal one would expect for $\theta_1 = 45°$ that the value of Δ would fall below 180° as the film thickness increased from zero. For convenience one may consider films which do not conform to this expectation, i.e. those which have a 'hump' in their Δ values, to possess markedly 'non-metallic' optical constants. For both angles of incidence it is only for wavelengths of 4358 Å and below that theory and experiment even approximately agree.



(*a*) (*b*)

Fig. 3.

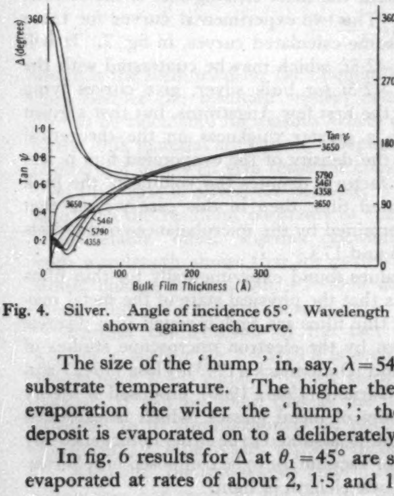


Fig. 4. Silver. Angle of incidence 65°. Wavelength shown against each curve.

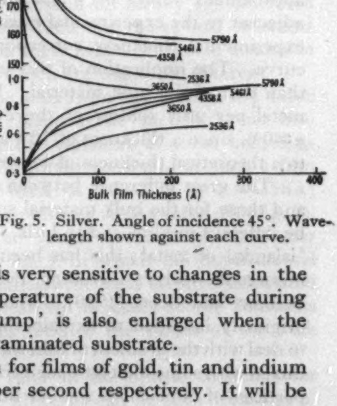


Fig. 5. Silver. Angle of incidence 45°. Wavelength shown against each curve.

The size of the 'hump' in, say, $\lambda = 5461$ Å is very sensitive to changes in the substrate temperature. The higher the temperature of the substrate during evaporation the wider the 'hump'; the 'hump' is also enlarged when the deposit is evaporated on to a deliberately contaminated substrate.

In fig. 6 results for Δ at $\theta_1 = 45°$ are shown for films of gold, tin and indium evaporated at rates of about 2, 1·5 and 1·5 Å per second respectively. It will be noted that the Δ values for the gold films all fall below 180°, in very approximate agreement with theoretical expectations. Both tin and indium, however, show large 'humps' which extend, especially for the latter, to far greater thicknesses than is the case for silver.

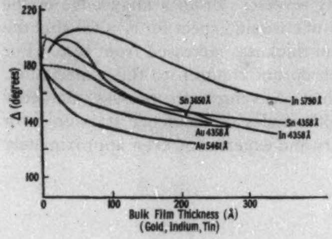


Fig. 6.

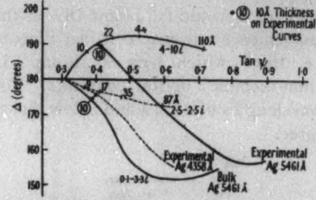


Fig. 7. Experimental and calculated curves at 45° incidence. Thickness (Å) shown by figures on curves.

§ 3. INTERPRETATION OF THE RESULTS

The direct determination of the optical constants of a thin film is tedious, for, it must be remembered, one cannot derive n and k, where the complex index of the film is $N=n-ik$, explicitly from measured values of ψ, Δ and t the film thickness, though the converse problem is easily soluble. The method adopted by the author was to calculate from the standard equations families of curves, with various parameters of n and k, for various values of t. The experimental curve was then fitted over a limited thickness range, by extrapolation if necessary, to one of the calculated curves, and thus an approximate value for the optical constants of the film was found.

If the results for silver are considered, the most striking fact is the sudden change between $\lambda = 5461$ Å and 4358 Å. The two experimental curves for these wavelengths are shown, together with some calculated curves, in fig. 7. It will be seen that constants $4-1{\cdot}0i$ and $2{\cdot}5-2{\cdot}5i$, which may be contrasted with the approximate values $0{\cdot}1-3{\cdot}3i$ and $0{\cdot}1-2{\cdot}3i$ for bulk silver, give curves lying adjacent to the experimental ones over the first few Ångströms, but that a given experimental thickness corresponds to a greater thickness on the theoretical curve. The implication of this is that the density of the evaporated film is less than that of the bulk material. If the factor q denotes the volume of the bulk metal per unit volume of the evaporated film, then in this case we see that $q \simeq 0{\cdot}6$, since a thickness of 10 Å as determined by the microbalance corresponds to a theoretical thickness of between 16 and 17 Å.

The great difference between the values found experimentally for thin films and those for the bulk material suggests that the physical state of the metal may be different in films and in bulk. Very thin films are agglomerated into discrete 'islands' of metal; this has been shown by the electron microscope studies of numerous workers including Picard and Duffendack (1943), Hass (1942) and Sennett and Scott (1950). Maxwell Garnett (1904, 1906) proposed a theory originally intended to explain the colours of metal glasses, which he extended to deal with the problem of thin metallic films, assumed to be either homogeneously mixed with some dielectric material (air or vacuum) or to be composed of spherica particles of dimensions smaller than the wavelength of light.

Maxwell Garnett's theory is not, perhaps, the most refined, but owing to the indeterminate physical nature of an evaporated film it is considered by the author that a simple theory is in many ways the best. Results similar to those of Garnett may be obtained by an extension of the well-known Lorenz–Lorentz equation. The additive properties of molecular refractivity defined by this equation may also be applied to mixtures if a proportionality factor W is included for each constituent i. For such a mixture we may write

$$\frac{\epsilon-1}{\epsilon+2}\frac{W}{S}=\sum_{1}^{i}\frac{\epsilon_i-1}{\epsilon_i+2}\frac{W_i}{S_i}, \qquad \ldots\ldots(1)$$

where ϵ is the dielectric constant and S is the density.

For a mixture of two components, both of unit permeability, eqn. (1) reduces to

$$\frac{n_{\mathrm{r}}^2-1}{n_{\mathrm{r}}^2+2}=\frac{n_{\mathrm{a}}^2-1}{n_{\mathrm{a}}^2+2}q_a+\frac{n_{\mathrm{b}}^2-1}{n_{\mathrm{b}}^2+2}q_{\mathrm{b}} \qquad \ldots\ldots(2)$$

where n_{r} is the resulting refractive index and q_{a} and q_{b} are the fractional volumes of the components a and b.

Rearranging eqn. (2) and substituting for $q_{\mathrm{b}}=1-q_{\mathrm{a}}$, we see that

$$\frac{n_{\mathrm{r}}^2-n_{\mathrm{b}}^2}{n_{\mathrm{r}}^2+2}=q_{\mathrm{a}}\frac{n_{\mathrm{a}}^2-n_{\mathrm{b}}^2}{n_{\mathrm{a}}^2+2}. \qquad \ldots\ldots(3)$$

The more complete theory of Maxwell Garnett leads to the following relation:

$$\frac{n_{\mathrm{r}}^2-n_{\mathrm{b}}^2}{n_{\mathrm{r}}^2+2n_{\mathrm{b}}^2}=q_{\mathrm{a}}\frac{n_{\mathrm{a}}^2-n_{\mathrm{b}}^2}{n_{\mathrm{a}}^2+2n_{\mathrm{b}}^2} \qquad \ldots\ldots(4)$$

which applies to a film composed of spherical particles.

If one considers a metallic film of index $n_{\mathrm{a}}=N=n-ik$ deposited *in vacuo*, then $n_{\mathrm{b}}=1$ and eqns. (3) and (4) both reduce to

$$\frac{n_{\mathrm{r}}^2-1}{n_{\mathrm{r}}^2+2}=q\frac{N^2-1}{N^2+2}, \qquad \ldots\ldots(5)$$

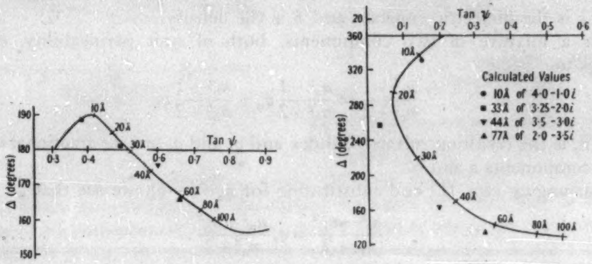
which is the form used for the calculations presented in this paper. The theory thus enables the resultant optical constants of such an ideally agglomerated film to be calculated for various values of q, provided that the optical constants of the bulk material are known. Depending on these initial constants, a suitable choice of q may well give resultant constants which are markedly different.

In the case of silver, for example, it can be seen from fig. 8 that for a value of $q\simeq 0{\cdot}6$ the resultant constants of the film change from 'metallic' ones into 'non-metallic' ones. Further calculation of the variation of these constants with wavelength shows that for $q\simeq 0{\cdot}65$ the resultant constants of such a film would change from about $4{\cdot}7-0{\cdot}8i$ at $\lambda=5461$ Å to $2{\cdot}5-3{\cdot}5i$ at $\lambda=4358$ Å; these values are similar to those found experimentally for such a value of q (fig. 7).

Further confirmation of the fact that an initial constant of the order $4-1{\cdot}0i$ is correct for $\lambda=5461$ Å is seen from fig. 9, in which the experimental values for $\theta_1=45°$ and $\theta_1=65°$ are shown together with some selected calculated points. It will be noted that as the film thickness increases so the constants required to fit simultaneously on both curves tend to the values of the bulk constants, i.e. n decreases whilst k increases.

Normally, the anomalous 'non-metallic' value of the optical constants for silver does not apply to any but the thinnest films, as can be seen from fig. 5.

This is presumably because the aggregates rapidly tend to coalesce and the constants thus tend towards those of the bulk material. If, for some reason, such as the raised temperature of the substrate, the aggregation persists to greater thicknesses, then the anomalous 'non-metallic' constants will similarly persist—this can be seen to be the case for the silver films deposited on a heated substrate, for the 'hump' in the Δ curve is much pronounced (fig. 10). Similarly, it may be argued that a slow evaporation rate or deposition on to a contaminated substrate favours the aggregated state and consequently makes the films more 'non-metallic' than normal.



(a) (b)

Fig. 9. Experimental results for silver films, 5461 Å.

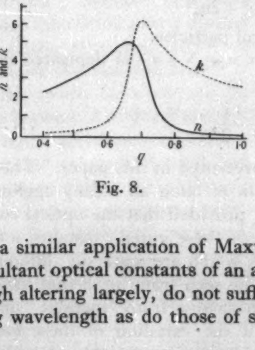


Fig. 8.

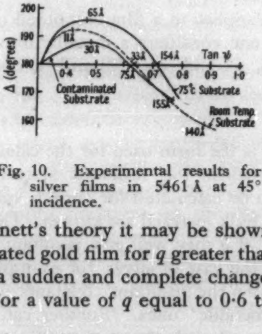


Fig. 10. Experimental results for silver films in 5461 Å at 45° incidence.

By a similar application of Maxwell Garnett's theory it may be shown that the resultant optical constants of an agglomerated gold film for q greater than 0·6, although altering largely, do not suffer such a sudden and complete change with varying wavelength as do those of silver. For a value of q equal to 0·6 table 1

Table 1. The Optical Constants of thin (10 Å) Gold Films ($q=0{\cdot}6$)

λ (Å)	5790	5461	4358
Approx. obs. values	$2{\cdot}5-3{\cdot}0i$	$1{\cdot}8-2{\cdot}5i$	$1{\cdot}6-1{\cdot}8i$
Theoretical values	$2{\cdot}3-3{\cdot}9i$	$1{\cdot}4-2{\cdot}4i$	$1{\cdot}6-1{\cdot}0i$
Bulk gold	$0{\cdot}23-2{\cdot}38i$	$0{\cdot}41-2{\cdot}05i$	$1{\cdot}29-1{\cdot}67i$

shows the experimental results together with those calculated theoretically, and, for comparison, the values of bulk gold as found by the author. As the 'hump' in the Δ curve for tin is so pronounced, agreement between experimental and Maxwell Garnett theoretical values was found up to thicknesses of 25 Å. Table 2 shows a comparison of results.

Table 2. The Optical Constants of thin (25 Å) Tin Films ($q = 0{\cdot}6$)

λ (Å)	5461	3560
Approx. obs. values	$3{\cdot}0-0{\cdot}5i$	$2{\cdot}3-1{\cdot}0i$
Theoretical values	$2{\cdot}8-0{\cdot}3i$	$2{\cdot}2-1{\cdot}5i$

It will be seen that there is in general good agreement between the experimental results and those predicted on the basis of an agglomerated film; as the real film is usually a complex system of aggregates and not very like the ideal film considered by Maxwell Garnett, it is somewhat surprising that the agreement is as good as it, in fact, is. The character of the results for tin and indium, being 'non-metallic' to fairly great thicknesses, suggests that these metals are aggregated up to thicknesses considerably greater than is the case for silver; indeed, electrical measurements on films of both metals deposited at room temperature showed that they were practically non-conducting, at least for thicknesses up to 300 Å. For thin films of tin some recent electron micrographs by Scott, McLaughlan and Sennett (1950) show the enhanced nature of the agglomerates very well.

§ 4. CONCLUSION

It has been shown that the application of Maxwell Garnett's theory of thin films is capable of explaining broadly the observed optical properties of evaporated metallic films. This implies that by far the most important factor which alters the optical constants of thin films from those of the bulk material is their aggregated structure. The permanence, as regards film thickness, of such a structure may be inferred from the observed optical properties of the film; this is made specially obvious when the bulk constants of the metal are such as to give 'non-metallic' constants for the thin aggregated films. The evaporation rate is found to be of importance.

.B
5
2

ACKNOWLEDGMENTS

The author wishes to acknowledge the advice and encouragement given him by Professor A. M. Taylor; he is also indebted to the Ministry of Education for a Further Education and Training grant.

REFERENCES

Archard, J. F., Clegg, P. L., and Taylor, A. M., 1952, *Proc. Phys. Soc.* B, **65**, 758.
Avery, D. G., 1949, *Nature, Lond.*, **163**, 916.
Clegg, P. L., and Crook, A. W., 1952, *J. Sci. Instrum.*, **29**, 201.
Drude, P., 1894, *Ann. Phys., Lpz.*, **51**, 77.
Hass, G., 1942, *Kolloid Z.*, **100**, 230.
Maxwell Garnett, J. C., 1904, *Phil. Trans. Roy. Soc.* A, **203**, 385; 1906, *Ibid.*, **205**, 237.
Picard, R. G., and Duffendack, C., O. S., 1943, *J. Appl. Phys.*, **14**, 291.
Rayleigh, Lord, 1902, *Scientific Papers*, **3**, 63 (Cambridge : University Press).
Scott, G. D., McLaughlan, T. A., and Sennett, R. S., 1950, *J. Appl. Phys.*, **21**, 843.
Sennett, R. S., and Scott, G. D., 1950, *J. Opt. Soc. Amer.*, **40**, 203.

Intensity Limitations in a Point Source of X-Rays

By V. E. COSSLETT
Cavendish Laboratory, Cambridge

MS. received 23*rd January, and resubmitted* 15*th May* 1952

ABSTRACT. When the diameter of the focal spot of electrons on an x-ray target is reduced to the order of a micron, the conditions of heat dissipation become so favourable that a greatly increased specific loading is possible. The use of an electron lens to obtain such a small spot introduces a limitation on the rate of supply of energy, as its high spherical aberration requires insertion of a very small aperture, the size of which must be reduced *pari passu* with the size of spot desired. For magnetic lenses of usual design the rate of energy supply is then proportional to the radius of the spot. Comparison with the rate of heat dissipation shows that, for spots of the order of a micron, the limitation on x-ray intensity is not the thermal properties of the target, as in tubes with large focus, but the emission obtainable from the cathode. Specific loadings of the order of 10^7 watts cm^{-2} can be tolerated in a spot of 1 micron; the corresponding emission current density would be of the order of 100 amp cm^{-2}.

---

§ 1. INTRODUCTION

In the normal type of x-ray tube, with a source of the order of 1 mm^2 in area, the main limitation on output intensity is the rate of heat dissipation in the target. As the size of the spot on the target is reduced the conditions of heat flow become more favourable (Müller 1927, 1929, 1931, Goldsztaub 1947, Oosterkamp 1948 a, De Barr and MacArthur 1950). Several types of microfocus tube have been made for purposes where high intensity rather than large total output is important (Ehrenberg and Spear 1951, Witty and Wood 1950). Special conditions obtain when a magnetic lens is used to obtain a very small spot and when the target is in the form of a Lenard window of thickness comparable with the range of the incident electrons, so that heat flow is almost entirely radial.

§ 2. HEAT DISSIPATION

When the radius of the electron spot r_2 is large compared with the target thickness t, heat flow can be regarded as purely axial from the front surface at temperature T_1 to the rear surface at T_2. If the thermal conductivity of the target material is k, the rate of dissipation of energy is

$$E_{\mathrm{x}} = \pi r^2 W_{\mathrm{x}} = \pi r_2^2 (T_1 - T_2) k/t, \qquad \ldots\ldots(1)$$

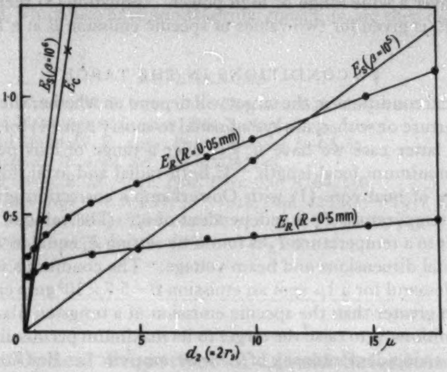
where W_{x} is the intensity or specific loading. When r_2 is of the same order as t radial heat flow becomes important. Oosterkamp (1948 b) obtains the combined energy flow E_{c} by introducing into eqn. (1) a correction factor s, dependent on the ratio of r_2 to t. For a constant value of s, as obtained by imposing the condition $r_2 = t$, we have $E_{\mathrm{c}} \propto r_2$ and $W_{\mathrm{c}} \propto 1/r_2$. The same simple relation holds when $r_2 \ll t$.

When the target is so thin that electrons penetrate through it, as in a microfocus tube recently constructed (Cosslett and Nixon 1952), energy is supplied to the metal throughout an approximately cylindrical volume. Assuming that this supply is uniform along the path t, and that the front and rear faces of the target are at the same temperature T_1, the flow of heat will be purely radial and

$$E_R = k(T_1 - T_2) 2\pi t / \ln(R/r_2), \qquad \ldots\ldots(2)$$

where R is the radius at which the target has the temperature T_2. In practice convection at the rear face of the target will lower its temperature below T_1; radiation in all practical conditions is negligible.

The figure shows the variation with spot diameter of E_c and E_R for a copper



Rate of supply and dissipation of energy for a target spot of diameter d_2 and a 10 kv electron beam. E_s=rate of supply for given value of β. E_R=rate of dissipation by radial heat flow for given target radius R. E_c=rate of dissipation by combined radial and axial flow.

target of thickness 1 μ. It is assumed that $T_2=100°\text{c}$ and $T_1=800°\text{c}$; the same results will obtain for a tungsten target at $T_1=2000°\text{c}$. The two curves for E_R assume values of $R=0{\cdot}05$ and $0{\cdot}5$ mm respectively.

§ 3. HEAT SUPPLY

As in optics, the energy flux in an electron lens system is constant across any plane normal to the axis, that is, β, the energy per unit area per unit solid angle, is invariant with axial distance for a given beam voltage and in the absence of absorption and scattering. If ρ is the current density at any plane at which the potential is ϕ and the angular width of the beam α, then $\beta=\rho/\pi\alpha^2$. If r_2 is the radius of the image, the power incident on it is

$$E_s=\pi r_2^{\,2}\rho_2\phi_2=\pi^2 r_2^{\,2}\alpha_2^{\,2}\beta\phi_2=\text{const.}, \qquad \ldots\ldots(3)$$

since $r_2\alpha_2$ is invariant, according to the electron optical form of the Helmholtz–Lagrange equation. Thus, if an electron lens were perfect, the rate of energy supply would remain the same whatever the size of the image spot formed, whereas the rate of dissipation in a metal target would fall with spot size according to eqn. (1) or eqn. (2), depending on the conditions.

In fact magnetic electron lenses suffer from severe spherical aberration, for which no correction is known. The aperture of the imaging beam has to be limited to make the radius of the disc of confusion δ_s less than that of the desired spot. If a lens has a spherical aberration coefficient C_s, then

$$r_2>\delta_s=C_s\alpha_2^{\,3}. \qquad \ldots\ldots(4)$$

In practice it is sufficient to prescribe $r_2=\delta_s$, as the disc of minimum confusion is smaller than that in the gaussian image plane. Substituting in eqn. (3),

$$E_s=\pi^2\beta r_2^{\,8/3}\,\phi_2 C_s^{\,-2/3}. \qquad \ldots\ldots(5)$$

C_s, which is of linear dimensions and of the order of the focal length f, is constant only at fixed power in a magnetic lens ; as the excitation is changed, C_2 varies rapidly. In the case of some simple field distributions it is found that $C_s \propto f^3$, but for actual lenses it follows from recent measurements (Liebmann and Grad 1951) that $C_s \propto f^{5/2}$ over a wide range of high power. Equation (5) then yields $E_s \propto r_2$; in the figure E_s is given for two values of specific emission β at a beam voltage of 10 kv.

§ 4. CONDITIONS IN THE TARGET

The thermal conditions in the target will depend on whether the lens is operated with fixed aperture or with α_2 always adjusted to satisfy eqn. (4) for the chosen spot size. In the latter case we have $E_s \propto r_2$ over a range of lens power practically down to the minimum focal length. If both radial and axial flow contribute to the dissipation of heat eqn. (1) with Oosterkamp's correction gives $E_c \propto r_2$ also. Hence target temperature T_1 is independent of r_2. The value of β which would raise the target to a temperature T_1 is found by setting E_s equal to E_c and inserting the experimental dimensions and beam voltage. The conditions which give E_c as in the figure demand for a 1μ spot an emission $\beta = 6{\cdot}7 \times 10^5$ amp cm^{-2} steradian^{-1}, which is much greater than the specific emission of a tungsten filament in normal operating conditions ; to raise the target to its maximum permissible temperature would require an incident intensity of $7{\cdot}5 \times 10^3$ amp cm^{-2}. For this to be supplied through the aperture fixed by eqn. (4) the source would have to emit 30 amp cm^{-2}. By overrunning the filament the emission could be raised towards the value needed to make optimum use of the thermal properties of the target, but at the cost of reducing the filament life to the order of a few hours. Increase in accelerating voltage ϕ_2, if permitted by other considerations, might be a preferable alternative (eqn. (3)).

Conditions become worse after the minimum focal length of the lens has been reached, at given ϕ_2. Further reduction in spot size can only be obtained by reducing the size of the electron source or increasing its distance from the lens. The rate of energy supply is then given by eqn. (5) with C_s constant, so that $E_s \propto r_2^{8/3}$ and diverges rapidly from the limiting rate of dissipation E_c if β is constant.

In practice a lens will often be operated with an aperture of fixed size, such that the spherical aberration is less than the minimum spot size required. Then in eqn. (3) α_2 is fixed and $E_s \propto r_2^2$. The rate of energy supply falls faster, with diminishing spot size, than does the rate of dissipation which, from eqn. (1) with Oosterkamp's correction, follows $E_c \propto r_2$. For a given maximum T_1 a condition of energy balance will be reached at a spot size depending on the limiting value of β. The tube will be most efficient in x-ray production if β can be made great enough to raise the target spot to the maximum permissible temperature at the smallest value of r_2 required.

As shown above, this is possible below $2r_2 = 1\mu$ only at the cost of shortened filament life. If it can be realized, then E_s will be much larger than E_c for appreciably larger target spots, and unless the filament emission is reduced when the spot size is increased the target may reach the melting point.

§ 5. CONCLUSIONS

A target spot smaller than about 1μ can only be raised to the maximum safe operating temperature by so increasing the intensity of the electron source that its lifetime is appreciably reduced. As the spot size is reduced a limit will be reached in which the lifetime of the filament becomes equal to the required exposure. The emissive power of the filament is thus the limiting factor on x-ray intensity and not the thermal properties of the target, as in a normal tube. The criterion in choosing a target metal then becomes efficiency of x-ray production, instead of thermal conductivity and melting point. Metals of high atomic number, such as tungsten and gold, which also have good thermal constants, are then to be preferred to copper.

REFERENCES

COSSLETT, V. E., and NIXON, W. C., 1952, *Proc. Roy. Soc.* B, in the press.
DE BARR, A. E., and MACARTHUR, I., 1950, *Brit. J. Appl. Phys.*, **1**, 305.
EHRENBERG, W., and SPEAR, W. E., 1951, *Proc. Phys. Soc.* B, **64**, 67.
GOLDSZTAUB, S., 1947, *C. R. Acad. Sci., Paris*, **224**, 458.
LIEBMANN, G., and GRAD, E. M., 1951, *Proc. Phys. Soc.* B, **64**, 956.
MÜLLER, A., 1927, *Proc. Roy. Soc.* A, **117**, 30 ; 1929, *Ibid.*, **125**, 507 ; 1931, *Ibid.*, **132**, 464.
OOSTERKAMP, W. J., 1948 a, *Philips Tech. Rep.*, **3**, 58 ; 1948 b, *Ibid.*, **3**, 303.
WITTY, R., and WOOD, P., 1950, *Brit. J. Appl. Phys.*, **1**, 307.

The Optical Properties of Liquid Selenium

BY E. W. SAKER

Physics Department, University of Reading

Communicated by R. W. Ditchburn; *MS. received* 8*th April* 1952, *and in final form* 14*th July* 1952

ABSTRACT. The optical properties of liquid selenium in the near infra-red have been measured for a temperature range up to about 400° c. The absorption edge moves out to longer wavelengths as the temperature increases by an amount corresponding to a change of energy gap of $1{\cdot}4\times10^{-3}$ ev/c°. Absorption is very small at low frequencies. For long wavelengths the refractive index decreases as the temperature increases by an amount corresponding to the thermal expansion.

It has been shown by many workers that the optical absorption of amorphous selenium falls rapidly with increasing wavelength in the red end of the visible spectrum and is very low in the infra-red. The photo-conductivity threshold does not coincide with the absorption edge (Weimer 1950, Weimer and Cope 1951, Gilleo 1951). Both the absorption edge and the photo-conductivity threshold move towards shorter wavelengths with decreasing temperature. As far as is known, the absorption of liquid selenium has not previously been measured. Since the structures of liquid and amorphous selenium are believed to be similar in nature, it is to be expected that liquid selenium will show optical properties very similar to those of amorphous selenium.

The absorption of liquid selenium as a function of wavelength was measured for a temperature range up to about 400° c. The selenium was contained in fused silica absorption cells which fitted closely into an electrically heated brass block. The thicknesses of cell used were 5 mm, 1 cm and 2 cm. The cells were placed in a parallel beam of light at the entrance slit of an infra-red monochromator. The light source was a tungsten filament lamp, the light from which was interrupted at a frequency of 800 c/s. The receptor cell was a lead sulphide photo-conductive cell which was coupled to a high-gain tuned amplifier. Figure 1 shows transmission against wavelength for a cell of 2 cm thickness. The movement of the absorption edge towards longer wavelengths as the temperature increases is very marked: in addition the edge (plotted in terms of wavelength) becomes less sharp. The absorption coefficients were obtained by the usual method of plotting log (transmission) against cell thickness. Figure 2 shows absorption coefficients plotted against frequency. The curve retains much the same shape over the whole of the temperature range. Using the present technique it was not possible to measure absorption coefficients greater than about 20 cm^{-1}. The fact that absorption is very small at low frequencies indicates that the free carrier concentration is small; this is confirmed by the fact that liquid selenium possesses a high thermopower (Borelius and Gullberg 1944).

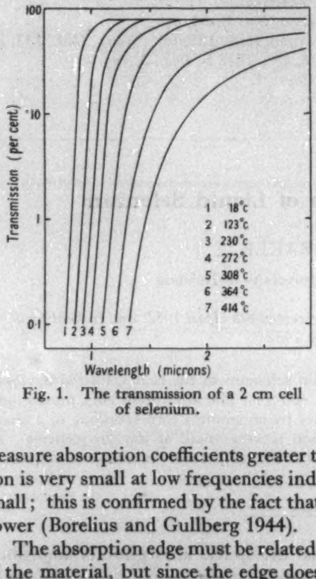


Fig. 1. The transmission of a 2 cm cell of selenium.

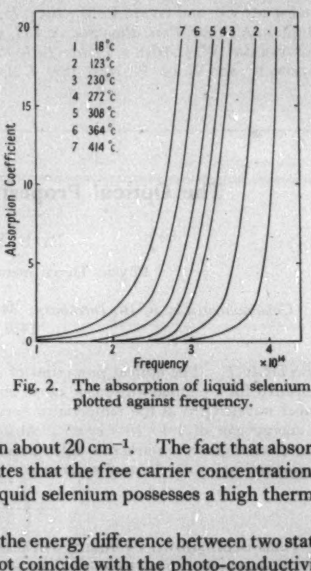


Fig. 2. The absorption of liquid selenium plotted against frequency.

The absorption edge must be related to the energy difference between two states of the material, but since the edge does not coincide with the photo-conductivity threshold this energy difference does not correspond in a perfectly direct way with a 'forbidden gap' width. Defining an energy $E_{20}=h\nu_{20}$ where ν_{20} is the frequency corresponding to an absorption coefficient of 20 cm^{-1}, it is found that $E_{20}=2{\cdot}16-1{\cdot}4\times10^{-3}T$ ev (T in °K). In a similar way $E_{1\cdot0}=1{\cdot}86-1{\cdot}4\times10^{-3}T$. Two determinations of the photo-conductivity threshold by Gilleo (1951) satisfy the equation $E_{\mathrm{P}}=2{\cdot}62-1{\cdot}4\times10^{-3}T$.

The refractive index of liquid selenium was measured by using a cell with faces inclined at an angle of about 20 degrees. The cell mounted in the heating block was placed on the table of a spectrometer of which the collimating slit coincided with the exit slit of the monochromator. The eyepiece of the spectrometer was replaced by the lead sulphide cell. The refractive index as a function of temperature and wavelength is shown in fig. 3. The values for the amorphous solid at 18° c agree to within 1% with previous measurements (Dowd 1951, Gebbie and Saker 1951). When the index is known the amount of radiation lost by reflection from the two silica–selenium and the two silica–air surfaces of the absorption cell can be calculated. The result of this calculation agrees to within a few per cent with a measurement of the amount of reflected radiation obtained by extrapolating to zero thickness a graph of log (transmission) against thickness. For long wavelengths where the dispersion is small the variation of $[n]$ with temperature (where $[n]=(n^2-1)/(n^2+2)$ and n is the refractive index) is such that $(1/[n])\,d[n]/dT=1{\cdot}7\times10^{-4}$ per °c. The writer has measured the expansion of solid and liquid amorphous selenium by interference and dilatometer methods respectively, and for both types of material the volume expansion coefficient is $1{\cdot}5\times10^{-4}$ per °c. The refractive index therefore appears to depend on the density ρ in the manner expected from theoretical considerations, so that $(n^2-1)/(n^2+2)\rho$ is constant with temperature.

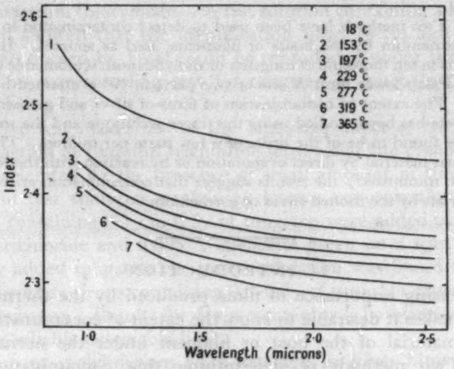


Fig. 3. The refractive index of liquid selenium.

ACKNOWLEDGMENTS

Thanks are due to Professor R. W. Ditchburn for the provision of research facilities and for much helpful advice, and to Messrs. Standard Telecommunication Laboratories Ltd. for a research grant.

REFERENCES

Borelius, G., and Gullberg, K., 1944, *Ark. Mat. Astr. Fys.* A, **31**, 1.
Dowd, J. J., 1951, *Proc. Phys. Soc.* B, **64**, 783.
Gebbie, H. A., and Saker, E. W., 1951, *Proc. Phys. Soc.* B, **64**, 360.
Gilleo, M. A., 1951, *J. Chem. Phys.*, **19**, 1291.
Weimer, P. K., 1950, *Phys. Rev.*, **79**, 171.
Weimer, P. K., and Cope, A. D., 1951, *R.C.A. Rev.*, **12**, 314.

The Contamination in Evaporated Films by the Material of the Source

By O. S. HEAVENS
Department of Physics, University of Reading

Communicated by R. W. Ditchburn; MS. received 23rd *March* 1952

ABSTRACT. Two methods have been used to detect contamination in evaporated films of silver and germanium by the boats or filaments used as sources. It is shown that a limit of a few parts in ten thousand of tungsten or molybdenum is attainable by microchemical methods. The much lower limit of one or two parts in 10^7 is attained by the radioactive tracer method. The extent of contamination of films of silver and germanium by tungsten and tantalum boats has been studied using the tracer technique and the minimum contamination attainable found to be of the order of a few parts per million. The possibilities of transfer of the boat material by direct evaporation or by reaction with the residual gas in the system have been minimized; the results suggest that contamination arises by the solution of the boat materials by the molten silver or germanium.

§1. INTRODUCTION

The increasing importance of films produced by the thermal evaporation process makes it desirable to know the extent of contamination of the films by the material of the boat or filament under the normal evaporation conditions. Two methods of determining this contamination have been studied: (i) microchemical analysis of the evaporate, (ii) use of radioactive tracer source elements. As would be expected, the latter method is considerably more sensitive than the former and is the only method capable of detecting contaminations smaller than 1 part in 10^6. The microchemical method can nevertheless provide a qualitative indication of the presence of impurities to the extent of a few parts in ten thousand. For cases in which this limit of detection is sufficiently low, the chemical method is to be preferred on account of its speed and simplicity.

The methods have been used to study the purity of films of silver and germanium when prepared by evaporation from boats or filaments of tungsten, molybdenum and tantalum. Two types of vacuum system have been used for these experiments. For tests at pressures down to 5×10^{-6} mm Hg, a metal system employing an oil diffusion pump was used; for pressures down to 10^{-7} mm Hg, a glass system with a mercury pump and solid CO_2 + acetone was used.

§2. MICROCHEMICAL TESTS

Using the methods given by Feigl (1947) films of silver prepared by evaporation (*a*) from a molybdenum boat and (*b*) from a tungsten filament were examined for contamination. The boats or filaments were first thoroughly cleaned and outgassed in vacuum at 2000° c and were then loaded with 'Specpure' silver. Films of area 2 cm^2 and thickness 3000–5000 Å were prepared at three evaporation temperatures, measured by optical pyrometer. The films were removed from the glass slide by the minimum necessary amount of aqua regia, on which the spot tests were made.

The thickness of the films was measured by placing a glass slide, partly covered with a razor blade, beside the target surface, thus enabling a portion of film with a well-defined edge to be collected. This slide was subsequently covered with an opaque silver layer and the step height determined using multiple-beam Fizeau fringes (Tolansky 1948). This method has been shown to be reliable for films of this order of thickness (Heavens 1951). The electron micrographs of Sennet and Scott (1950) show that the density of films of this thickness is unlikely to differ appreciably from that of the bulk material.

For the detection of molybdenum use is made of the brick-red coloration which results from the addition of stannous chloride to a mixture of a molybdate solution and potassium thiocyanate. A test solution containing a known amount of molybdenum showed that the identification limit ($0{\cdot}1\gamma$) and concentration limit (1 : 500 000) quoted by Feigl could be reproduced under the conditions of these tests. The corresponding test for tungsten is much less sensitive, the detection limit being 5γ.

An alternative test makes use of the catalysing effect of tungstates on the decolorizing reaction of titanium trichloride with malachite green. This action is markedly accelerated by the presence of small amounts of tungstates in acid solution. From test solutions containing known concentrations of tungsten, weighed drops containing $0{\cdot}1\gamma$ to $0{\cdot}5\gamma$ of tungsten were added to microdrops of 1% titanium trichloride and 0·005% malachite green on a spot plate. In the absence of any added tungstate the green coloration was found to disappear in from 3 to 10 minutes: drops containing $0{\cdot}2$–$0{\cdot}3\gamma$ of tungsten caused decoloration in from 1 to 2 minutes. Results for drops containing $0{\cdot}1\gamma$ were erratic, indicating $0{\cdot}2$–$0{\cdot}3\gamma$ as the reliable detectable limit.

No tungsten or molybdenum were detected in silver films for which the evaporation temperature was less than 1700° c. For the temperature range 1750–1800° c, just detectable amounts of the boat material were observed with some films. Taking $0{\cdot}2\gamma$ as the detection limit in each case the observed contamination at 1750–1800° c corresponds to 2–3 parts in 10^4. This figure is confirmed by the results of the radioactive tracer tests described below.

§ 3. RADIOACTIVE TRACER METHOD

Many properties of evaporated films are likely to be only negligibly affected by smaller amounts of contamination than are detectable by the microchemical methods discussed above. It is known, however, that the electrical properties of semiconducting materials are very sensitive to minute amounts of impurity, concentrations far below the limit detectable by chemical methods being sufficient to cause significant changes. For certain types of impurity, the use of the radioactive tracer technique enables the detection limit to be lowered by two or three orders of magnitude below the chemical limit. The availability of the method depends on the existence of an active isotope which has a suitable half-life and which can be produced in sufficient concentration. Measurements have been made using tungsten and tantalum, for which elements very suitable active isotopes may be produced by (n, γ) reactions. The tantalum isotope ^{182}Ta has a half-life of 117 days: since the whole of natural tantalum consists of ^{181}Ta, a large activity (80 mc/g) is obtainable. The detection limit attained with this material is less than 10^{-9} g. With the tungsten isotope ^{187}W, of which the half-life is 24·1 hours, an activity of 360 mc/g is obtained, yielding a detection limit of less than 10^{-10} g. For molybdenum, the tracer method offers but little advantage over the chemical method. The isotope ^{99}Mo possesses a suitable half-life (67 hours) but the low activity at present attainable enables a detection limit of only 10^{-8} g to be reached.

Small boats, made from 0·2 mm sheet, of tungsten and tantalum were irradiated in the Harwell pile until sufficient activity had resulted. The boats were then used as evaporation sources and the evaporated materials tested for active elements in the usual way.

§ 4. EXPERIMENTAL DETAILS

(i) *The Metal System*

The metal system is a standard vacuum evaporation plant with Pyrex bell-jar. By means of suitable long-handled tools, the active boat is installed and clamped. A platinum–platinum-rhodium thermocouple is mounted on a light spring so that when the boat is installed, the couple is held against the bottom of the boat. (The activity of the boat makes the usual spot-welding process of attachment too hazardous.) An optical pyrometer is used for temperatures above 1 600° c. A shutter is moved over the boat to collect the first part of the evaporate which may consist of surface contamination, for example from oxidation by the atmosphere. All vacuum taps, shutters etc. are handled by remote control.

(ii) *The Glass System*

The glass system consists of a large U-tube mounted horizontally (fig. 1). The tungsten boat is mounted on the end of two tungsten rods, 3 mm diameter and 60 cm long, which are prevented from sagging by glass bridges. The targets are nickel discs of the type used in the standard Geiger counter assembly. In the glass system these are held on a glass spoon which is kept out of sight of the boat during preliminary outgassing. A magnet enables the target to be moved into position to receive the evaporated material.

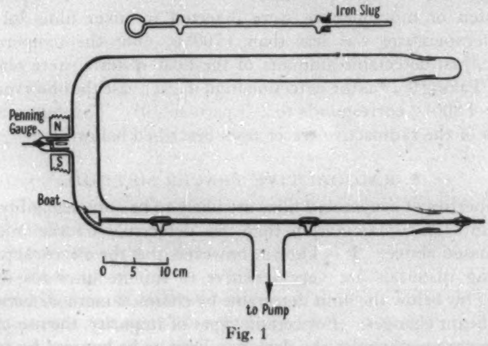


Fig. 1

Determination of the sensitivity.

A minute (~100γ) weighed piece of boat material is attached to aluminium foil with collodion and irradiated in the pile simultaneously with the boat. An identical aluminium foil plus collodion is also irradiated so that the (small) activity due to these supporting materials may be readily allowed for. (This activity was found to decay to negligible proportions in a few hours.) The use of a small weighed specimen, for which the disintegration rate is within reach of the normal counting apparatus enables the specific activity of the boats to be found easily.

The sensitivity obtained with the tantalum was such that N counts per minute above background corresponded to $2{\cdot}3_4 \times 10^{-10} \times N$ g. The slow decay of the ^{187}Ta isotope allows the fall in sensitivity over a period of a few days to be neglected. With tungsten the initial sensitivity was such that N counts per minute corresponded to $9{\cdot}5 \times 10^{-12} \times N$ g.

§ 5. RESULTS OF TESTS

(i) *Tests with Boats Alone*

A preliminary series of tests was made in which the boats, after initial cleaning and outgassing for several minutes at over 2000° C, were heated empty to various temperatures from 1000° C to 1900° C and the targets tested for radioactive material. The amounts of tungsten and tantalum transferred to the target were found to be considerably larger than would be expected from direct evaporation at the temperatures used. The mass of metal transferred was found to depend directly on the pressure in the system, results obtained for tungsten at 1200° C being shown in fig. 2. Substantially similar results were obtained with the tantalum boat, the rate of loss of tantalum ranging from 5×10^{-8} g cm^{-2} sec^{-1} at 5×10^{-6} mm Hg, to 45×10^{-8} g cm^{-2} sec^{-1} at 4×10^{-5} mm Hg. Over the range of pressures used, the rate of loss of boat material is directly proportional to pressure, suggesting that the mechanism of transfer involves reaction with the residual gas in the system.

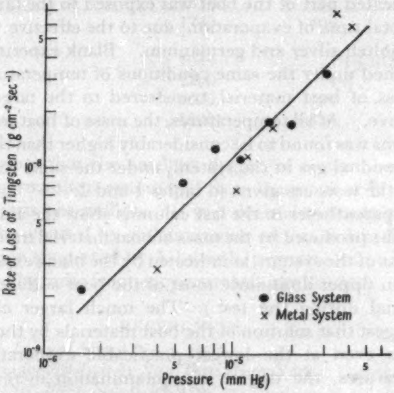


Fig. 2

The two most likely mechanisms would seem to be (i) oxidation of the boat by residual oxygen, (ii) oxidation by reaction with residual water vapour. The oxides of tungsten and tantalum thus produced evaporate immediately at temperatures above 1000° C. Langmuir (1915) has studied the first-mentioned process in the case of tungsten and finds that only a fraction of the incident oxygen atoms reacts with the tungsten. At a pressure of 10^{-5} mm and temperature of 1 200° c, Langmuir's figures lead to a rate of loss of tungsten of 5×10^{-10} g cm^{-2} sec^{-1} for this process, compared with the observed value of $1{\cdot}2 \times 10^{-8}$ g cm^{-2} sec^{-1} if it be assumed that the residual gas is air.

It is thus clear that reaction with residual oxygen will not account for the observed transfer of boat material. Moreover, observations on the nature of residual gases in vacuum systems of this type show that the predominant constituent is usually water vapour (Blears 1951). Tungsten is known to react readily with water vapour at the temperatures of these experiments (Langmuir 1913) so that some transfer of material by this process is to be expected. If it be assumed that the residual gas is entirely water vapour, then the number of water molecules striking the boat per second is sufficient to account for the observed transfer. From the point of view of studying the contamination of germanium and silver films, it is seen that, provided the pressure in the system is kept below 10^{-6} mm, no direct transfer of tungsten occurs within the limits of detection at an evaporation temperature of 1 200° c. Any tungsten detected in films produced at these pressures must therefore be carried up with the evaporated material, and further reduction of pressure in the system is unlikely to produce films of lower tungsten content.

(ii) *Tests with Silver and Germanium*

Weighed quantities of these materials were evaporated completely from the active boats and the mass of the films produced determined either directly by weighing or interferometrically. The silver or germanium charge was first melted so that the whole of the surface of the boat was covered. During the last stages of the evaporation, part of the boat surface became exposed. The time for which any of the heated part of the boat was exposed to the target did not exceed about 5% of the total time of evaporation, due to the effective wetting of the boat materials by the molten silver and germanium. Blank experiments using empty boats were performed under the same conditions of temperature and pressure to determine the mass of boat material transferred to the target by the process described in (i) above. At all temperatures, the mass of boat material in the silver and germanium films was found to be considerably higher than the mass transferred by reaction with residual gas in the system, under the same conditions.

The results of the tests are given in tables 1 and 2.

The figures in parentheses in the last columns show the amount of contamination which would be produced by the mass of boat material transferred by reaction with the residual gas of the system, as indicated by the blank experiments. (These figures represent an upper limit since most of the boat surface is covered by the evaporating material during the test.) The much larger observed values of contamination suggest that solution of the boat materials by the molten silver and germanium occurs even at the lowest practicable evaporation temperatures. At higher temperatures, the degree of contamination increases considerably; silver films containing over 10% of tantalum have been produced by evaporating at temperatures of about 2 500° c.

In some experiments at the lower temperatures (1000° c–1 200° c), the evaporation was stopped before any of the boat surface became exposed. The figures obtained were found not to differ significantly from those in which the boat contents were evaporated completely and have been included in the mean values for each temperature which are given in tables 1 and 2.

Table 1. Tests in Metal System (pressure 8–15 × 10^{-6} mm)

Boat Material	Evap. Metal	Evap. Temp. (°c)	Contamination of Film × 10^{6}	
W	Ag	1000	4·9 ± 0·3	(< 0·5)
		1100	22	(1)
		1410	120	(20)
		1900	330	(90)
	Ge	1070	2·3 ± 0·3	(< 0·5)
		1250	5·1	(1·5)
		1470	64	(27)
		1800	430	(105)
Ta	Ag	1100	20	(3)
		1200	29	(3)
		1410	45	(10)
		1670	113	(29)
	Ge	1120	11	(3)
		1380	7	(4)
		1530	17	(12)
		1800	76	(37)

Table 2. Tests in Glass System (pressure below 5 × 10^{-7} mm Hg)
Tungsten Boat

Evap. Metal	Evap. Temp. (°c)	Contamination of Film × 10^{6}	
Ag	1080	12 ± 2	(0)
	1200	34	(0)
	1510	68	(6)
	1740	106	(19)
Ge	1040	2·1 ± 0·2	(0)
	1200	4·3	(0)
	1620	39	(10)
	1800	330	(42)

It appears therefore that the smallest degree of contamination attainable in films of silver and germanium evaporated from tungsten or tantalum boats is of the order of a few parts per million. If it is desired to study properties of films which are likely to be affected by contamination of this order, then an alternative method of evaporation suggested by Ashworth (private communication) in which the heating is effected by radio frequency currents may be used.

REFERENCES

BLEARS, J., 1951, *J. Sci. Instrum*, Suppl. No. 1, 36.

FEIGL, F., 1947, *Qualitative Analysis by Spot Tests* (New York : Elsevier).

HEAVENS, O. S., 1951, *Proc. Phys. Soc.* B, **64,** 419.

LANGMUIR, I., 1913, *Trans. Instn. Elect. Engrs.*, 32, 1893 ; 1915, *J. Amer. Chem. Soc.*, **37,** 417.

SENNET, R. S., and SCOTT, G. D., 1950, *J. Opt. Soc. Amer.*, **40,** 203.

TOLANSKY, S., 1948, *Multiple Beam Interferometry* (Oxford : University Press).

Self-Absorption in Arc Sources in Thermal Equilibrium

By H. EDELS
Department of Electrical Engineering, University of Liverpool

MS. received 8*th May* 1952

ABSTRACT. An analytical determination of the self-absorption of a spectral line emitted by a source is not normally possible if account is taken of the spatial variations in the source of the excited atom concentrations and spectral line profiles. The problem is simpler for uniformly excited sources, for example sources in thermal equilibrium, but not for high-pressure arc sources normally assumed to be in thermal equilibrium, because of the existence of temperature gradients. In the paper an analysis is developed which shows that if such arcs are treated as uniformly excited, an error of less than 5% is introduced into the determination of the self-absorption of spectral lines, emitted by transitions between an upper state s and a lower absorbing state t, which satisfy the limits $h\nu_{sq}/kT_0 \geqslant 10$ and $\nu_{st}/\nu_{tq} \leqslant 0{\cdot}5$, where q is the ground state and T_0 the axial source temperature. The theory is then applied to the experimental data for a high-pressure mercury arc discharge and to new measurements of the line widths at 5461 and 4358 Å. It is shown that the calculated value of the self-absorption at 5461 Å agrees with the experimental value only when the line has a resonance profile. The existence of this form of profile is also indicated by considerations of the broadening processes in the discharge.

§1. INTRODUCTION

When radiation of a given frequency is emitted by any small volume of a gaseous light source it has normally to pass through a portion of the source before being received by the observer. During this process a fraction of the emitted radiation energy is absorbed by atoms and molecules similar to those causing the emission. This phenomenon is termed self-absorption, and a knowledge of its magnitude is often required in spectral investigations of gaseous electrical discharges.

In order to determine the self-absorption account must be taken of the spatial variations of excited atom concentrations and spectral line profiles. The complexity of the problem is such that analytical determinations are not normally possible, and recourse must be made to graphical methods. Such methods are, however, laborious, so that simplified source conditions are normally assumed in order to obtain analytic solutions. This is evident in the comprehensive work of Cowan and Dieke (1948), who obtain the general equation and consider its application to the characteristics of different source models and the extent to which these models are comparable with a practical source. In the exhaustive treatment of Bartels (1949, 1950), although a rigorous solution of the problem is not achieved, an approximate general solution is obtained which is applied to the special case of self-reversal, when the reversal peaks have reached their maximum value.

From the work of these and previous investigators it is clear that the mathematical difficulties arise from the condition that the concentration and line profile variations of the emitting and absorbing states are different, and that the problem is considerably simplified for uniformly excited sources, that is, sources throughout which the ratio of the concentrations of emitting and absorbing atoms is constant. This condition holds, for example, for sources in thermal equilibrium, in which the energy of all types of particle are distributed according to the Maxwell–Boltzmann law. However, in arc sources which are accepted as being in a state of thermal equilibrium, for example in the carbon arc in air, as shown by Ornstein and Brinkman (1934) and Kruithof and Smit (1944), and the mercury arc by Elenbaas (1951) (see also Edels 1950), appreciable temperature gradients exist, as shown, for example, by Kruithof (1943), Fischer and König (1938), Elenbaas (1951), so that the assumption of thermal equilibrium is invalid when applied to the whole discharge region. The following analysis of self-absorption shows that, despite the existence of temperature gradients, it is often possible to assume that the source is uniformly excited without causing undue errors in the final results. This makes possible self-absorption calculations of the more complex cases of these sources, including line profile variation. Such calculations are of considerable importance because the presence of self-absorption in a source in thermal equilibrium produces emergent radiation which is not descriptive of a thermal source.

§ 2. DERIVATION OF A SIMPLIFIED SELF-ABSORPTION EQUATION

We will first obtain the general equation for self-absorption and derive from it the simplified equation which holds for the special case of sources in thermal equilibrium.

Consider a light source in a steady state so that the emission and absorption processes have no effect on the concentrations of the atoms in the excited states. If we then take a small volume dv of the source containing dN_s atoms/cm³ in an upper state s capable of emitting radiation by transition to a lower state t with a frequency between ν and $\nu+d\nu$, and dN_t atoms/cm³ in the lower state capable of absorbing this radiation, and if this volume is subject to an intensity of radiation I_ν, then the energy received from the solid angle dw is

$$\Delta(E_\nu d\nu)=\{dN_sA_{st}-(dN_tB_{ts}-dN_sB_{st})I_\nu\}h\nu\,dw\,dv/4\pi \qquad \ldots\ldots(1)$$

where A_{st}, $(B_{st}I_\nu)$ and $(B_{ts}I_\nu)$ are Einstein's probabilities of spontaneous and induced emission and absorption respectively. Now in most investigations the solid angle of the observed radiation from the source is small, and the radiation may be taken to exist as a plane wave directed along the axis of observation x. Thus the change of intensity over dx is

$$d(I_\nu d\nu)=\left\{dN_sA_{st}\frac{h\nu}{4\pi}-\frac{dN_tA_{st}\lambda^2g_s}{8\pi g_t}\left(1-\frac{dN_sg_t}{dN_tg_s}\right)I_\nu\right\}dx \qquad \ldots\ldots(2)$$

where the relationships between the Einstein coefficients have been substituted and g_s and g are the statistical weights of the excited states. In most practical cases Kirchhoff's law is obeyed, so that the spectral absorption and emission frequency distributions are the same at any point. Thus, assuming a symmetric source about $x=0$ such that $(N_s)_x=(N_s)_0f(x)$ and $(N_t)_x=(N_t)_0F(x)$, we have

$$(dN_s)_x=(N_s)_0f(x)f(\nu,x)\,d\nu \quad \text{and} \quad (dN_t)_x=(N_t)_0F(x)f(\nu,x)\,d\nu, \qquad \ldots\ldots(3)$$

where $(N_s)_0$ and $(N_t)_0$ are the total concentrations of excited atoms in the states s and t at $x=0$, and a normalization condition exists given by $\int f(\nu,x)\,d\nu=1$ over the range $-\infty$ to $+\infty$. Substituting (3) into eqn. (2) we obtain the general self-absorption equation

$$d(I_\nu d\nu)=\{(I_e)_0f(x)f(\nu,x)\,d\nu-K_0F(x)f(\nu,x)(I_\nu d\nu)\}dx \qquad \ldots\ldots(4)$$

where $(I_e)_0=(N_s)_0A_{st}h\nu/4\pi$, $K_0=\{(N_t)_0A_{st}\lambda^2g_s/8\pi g_t\}\{1-(N_s/N_t)_xg_t/g_s\}$.

$(I_e)_0$ is the true intensity emitted at $x=0$ and K_0 is the integral of the absorption coefficient over the whole spectral line at $x=0$, as given, for example, by Mitchell and Zemansky (1934). The term $1-(N_s/N_t)_x g_t/g_s$ is in most practical cases nearly unity and may be ignored, although with the approximation which is to be made it may be reintroduced later.

If the variations along the x axis are assumed to be even functions and we take a boundary for the line of frequency ν of $\pm X$, then the solution of eqn. (4) gives the received intensity in the frequency interval $d\nu$ as

$$I_\nu\, d\nu=(I_e)_0\, d\nu\{\exp[-K_0(\psi)_0{}^X]\}\int_{(\psi)^0}^{(\psi)^X}\frac{f(x)}{F(x)}\{\exp(K_0\psi)+\exp[K_0(2(\psi)^0-\psi)]\}\, d\psi \qquad \ldots\ldots(5)$$

where $(\psi)^x=\int^x F(x)f(\nu,x)\,dx$. With the general solution in this form it will be apparent that the difficulties in applying it to practical cases are due almost entirely to the presence of the factor $\{f(x)/F(x)\}$ and that the expression is appreciably simplified for a uniformly excited source since $f(x)/F(x)$ is a constant. If we now consider a source which has a temperature gradient of such a magnitude that the source still approximates to a state of thermal equilibrium, then it has been shown (Edels 1951) that $\{F(x)\}^\beta=f(x)$ if $\boldsymbol{h}\nu_{sq}/\boldsymbol{k}T_0\geqslant 10$ and $\nu_{st}/\nu_{tq}\leqslant 0{\cdot}5$ where $\beta=\nu_{sq}/\nu_{tq}$, where q represents the ground state and T_0 is the temperature at the source centre. The limit of ν_{st}/ν_{tq} precludes the application of this relationship to resonance lines, a restriction which appears also in Bartels' work. With $\{F(x)\}^\beta=f(x)$ in eqn. (5) the solutions are still laborious. However, for many spectral lines β does not differ greatly from unity, thus the Balmer series gives $\beta_{\max}\sim 1{\cdot}3$, and since it is the integrated values of $F(x)$ and $f(x)$ which affect absorption calculations, it would appear feasible that no serious error will result by putting $\beta=1$. That such an approximation is possible is shown by the following calculations.

To simplify the calculations assume that the spectral line profile does not vary spatially and consider only that half of the source furthest from the observer, with $x=0$ at the furthest point. Thus with $\{F(x)\}^\beta=f(x)$ we have

$$I_\nu\, d\nu=(I_e)_0\, d\nu f(\nu)\exp\left[-K_0 f(\nu)\int_0^X F(x)\,dx\right]\int_0^X\{F(x)\}^\beta\exp\left[K_0 f(\nu)\int_0^x F(x)\,dx\right]dx. \qquad \ldots\ldots(6)$$

If we now take a function $F(x)=(x/X)^n$ and assume a linear spectral profile of half-width $\Delta\nu$ and central frequency ν_0 given by $f(\nu)=(1/\Delta\nu)\{1-|(\nu-\nu_0)/\Delta\nu|\}$, then from (6) the total received intensity I, as a fraction of the true emitted intensity I_e which would be received if no absorption were taking place, is given by

$$P=\frac{I}{I_e}=\left[1-2\left\{\frac{b}{3(\phi+2)}-\frac{b^2}{4(\phi+2)(\phi+3)}+\frac{b^3}{5(\phi+2)(\phi+3)(\phi+4)}-\ldots\right\}\right] \qquad \ldots\ldots(7)$$

where $b=XK_0/\Delta\nu(n+1)$ and $\phi=n(\beta-1)/(n+1)$. When $\beta=1$, $\phi=0$, and the fraction of the intensity received is given by

$$Q=\frac{I}{I_e}=\left[1-2\left\{\frac{b}{3.2!}-\frac{b^2}{4.3!}+\frac{b^3}{5.4!}-\ldots\right\}\right]. \qquad \ldots\ldots(8)$$

The maximum error in assuming $\beta=1$ for a given value of b is obtained when ϕ is a maximum; for hydrogen this occurs when $\beta=1{\cdot}3$, so that with $n=\frac{1}{2}$,

$\phi = 0{\cdot}1$. In this case when $b=1$, $Q=0{\cdot}736$ and $P=0{\cdot}746$, giving an error of slightly greater than 1% when the absorption taking place is of the order of 26%. When $b=2$ the absorption is 43% and the error is only 2·3%. Now for the maximum value of ν_{st}/ν_{tq} assumed previously, the corresponding maximum for β is 1·5. With this value and $b=1$ and 2 we obtain $P=0{\cdot}753$ and 0·593 respectively, giving errors less than 5%, which are within the range of accuracy normally obtained in intensity measurements. Thus for a source in thermal equilibrium satisfying the conditions $\boldsymbol{h}\nu_{sq}/\boldsymbol{k}T_0 \geqslant 10$ and $\nu_{st}/\nu_{tq} \leqslant 0{\cdot}5$, self-absorption calculations may be made assuming that $F(x)=f(x)$. With this assumption eqn. (5) reduces to

$$I_\nu \, d\nu = \frac{(I_e)_0}{K_0}[1-\exp\{-2K_0(\psi)_0{}^X\}]d\nu. \qquad \ldots\ldots(9)$$

This is the well-known equation for a source of constant excitation having a uniform emission $(I_e)_0$, uniform coefficient K_0 and an equivalent absorbing depth of $2\Delta\nu_0(\psi)_0{}^X$. With this equivalent depth we may utilize relationships previously obtained for constant excitation sources. Thus for a source of circular cross section of radius R, if we neglect spatial variation of spectral profile and consider the absorption along the source diameter we obtain the ratio of the received to true emitted intensity I/I_e for different profiles as follows:

Resonance:
$$f(\nu) = \frac{2}{\pi\Delta\nu_0}\bigg/\left[1+\left\{\frac{2}{\Delta\nu_0}(\nu-\nu_0)\right\}^2\right].$$

Then
$$I/I_e = \exp(-ap)[J_0(iap)-iJ_1(iap)] \qquad \ldots\ldots(10)$$

(Ladenburg and Reiche 1913) where $a=(2/\pi)\int_0^1 F(z)\,dz$; $z=r/R$; $p=RK_0/\Delta\nu_0$ and J_0, J_1 are Bessel functions of the first kind.

Doppler:
$$f(\nu) = \frac{2}{\Delta\nu_0}\left(\frac{\ln 2}{\pi}\right)^{1/2}\exp\left[-\ln 2\left\{\frac{2}{\Delta\nu_0}(\nu-\nu_0)\right\}^2\right].$$

Then
$$\frac{I}{I_e} = \left[1-\frac{ap}{2!\sqrt{2}}+\frac{(ap)^2}{3!\sqrt{3}}-\ldots\right] \qquad \ldots\ldots(11)$$

(Ladenburg 1930) where $a=4\{(\ln 2)/\pi\}^{1/2}\int_0^1 F(z)\,dz$ and $p=RK_0/\Delta\nu_0$.

Linear:
$$f(\nu) = \frac{1}{\Delta\nu_0}\left[1-\left|\frac{\nu-\nu_0}{\Delta\nu_0}\right|\right].$$

Then
$$\frac{I}{I_e} = \frac{2}{ap}\left[1-\frac{1}{ap}\{1-\exp(-ap)\}\right] \qquad \ldots\ldots(12)$$

where $a=2\int_0^1 F(z)\,dz$ and $p=RK_0/\Delta\nu_0$, by integration of (9). In each case the absorption is dependent upon p, the absorption parameter, and in general, for any profile varying spatially only in half-width, it can be shown that the absorption is always a function of p. Thus, expanding and integrating (9) we obtain the total received intensity of a spectral line from ν_0 to $\nu_0+\delta\nu$ as

$$I = 4(I_e)_0\int_{\nu_0}^{\nu_0+\delta\nu}\left\{\psi-\frac{2K_0}{2!}\psi^2+\frac{(2K_0)^2}{3!}\psi^3-\ldots\right\}d\nu.$$

The first term of this expression is the true emitted intensity I_e, so that

$$\frac{I}{I_e} = \left[1-\frac{2K_0}{2!}\frac{\int\psi^2\,d\nu}{\int\psi\,d\nu}+\ldots+\left(-\frac{2K_0}{S!}\right)^{S-1}\frac{\int_{\nu_0}^{\nu_0+\delta\nu}\left\{\int_0^R F(x)f(\nu,x)\,dx\right\}^S d\nu}{\int_{\nu_0}^{\nu_0+\delta\nu}\int_0^R F(x)f(\nu,x)\,dx\,d\nu}\right]. \qquad \ldots\ldots(13)$$

Now in general $f(\nu, x)=(c/\Delta\nu_x)f[(\nu-\nu_0)/\Delta\nu_x]$, and $\Delta\nu_x$ in terms of $\Delta\nu_0$ the half-width at $x=0$ is given by $\Delta\nu_x=\Delta\nu_0 f(x/R)$. Since $F(x)$ can also be stated as a function of x/R, substitution in (13) and change of variables gives for the general term $[\alpha_{S-1}(-RK_0/\Delta\nu_0)^{S-1}]$ where α is a constant. Thus (13) becomes

$$\frac{I}{I_e}=\left[1-\alpha_1\frac{RK_0}{\Delta\nu_0}+\alpha_2\left(\frac{RK_0}{\Delta\nu_0}\right)^2-\ldots.\right]=[1-f(p)].$$

The absorption parameter, after substitution of the expression for K_0, and giving the state concentrations their thermal values, is given by

$$p=[R(N_t)_0A_{st}\lambda^2g_s/8\pi g_t\Delta\nu_0][1-\exp(-\boldsymbol{h}\nu_{st}/\boldsymbol{k}T_0)].$$

From this it will be seen that an increase in absorption will take place with increases in the absorbing state concentration, source dimensions, transition probability and wavelength, but decreases with increase in half-width.

It is of interest to note that the analysis precludes self-reversal, since differentiation of eqn. (9) shows that the maximum of the received intensity distribution still occurs at the central frequency.

§ 3. SELF-ABSORPTION IN ARC SOURCES

The simplified self-absorption equation (9) may be applied only to those sources in which thermal equilibrium exists and to which the given limits apply, so that $F(x)$ may be taken to equal $f(x)$.

Since most high-pressure arc columns (~ 1 atm) are believed to be in thermal equilibrium, it should be possible therefore to obtain further confirmation of the validity of the simplified equation from the experimental data for this type of source. The data available are, however, extremely meagre: only the carbon and mercury arcs have been shown to be in thermal equilibrium, and self-absorption data have been obtained only for the latter case. We will use the data provided by Elenbaas (1951) for his mercury arc discharge (0·88 atm, 6 amp, $T_{gas}=5\,885°$ K at $x=0$), in which it is clear that the lines at 5461 and 4358 Å are self-absorbed. Dr. Elenbaas has kindly measured the half-width values of these lines and obtains 0·25 Å for 5461 Å and approximately the same for 4358 Å, so that only the profile shapes remain to be determined. A good indication of these may be obtained from a consideration of the processes affecting the line broadening in the discharge, and from the work of other investigators.

The natural width of the lines may be neglected as well as the Doppler width, which is approximately 0·02 Å at $T\sim 5\,900°$ K. Since neither Wendt and Wetzel (1916) nor Haslam (1935) obtained measurable Stark displacements of these lines, we assume that the quadratic Stark effect constants are less than 10^{-14}. With this value and an electron concentration of $5\times 10^{15}/\text{cm}^3$ obtained from Saha's equation with a reduced ionization potential of 9·7 ev (Schulz 1948), we obtain, using Lindholm's (1945) formula, a negligible Stark width of the order of 0·02 Å. It would appear therefore that the lines are mainly pressure broadened, and this is confirmed by the work of Rompe and Schulz (1938) and Schulz (1938) on mercury arcs at 20 atm. These investigators find that asymmetric pressure broadening and resonance profiles are associated with the series $(n^3S_1-2^3P_2)$, which includes the line at 5461 Å. It may be concluded therefore that the lines at 5461 and 4358 Å from a mercury arc at a pressure of approximately 0·9 atm also have a resonance profile. In this case the absorption of the lines may be determined from eqn. (10) if the limits of the analysis are satisfied. Calculation shows that for 5461 Å, $h\nu_{sq}/kT_0 \sim 15$ and $\nu_{st}/\nu_{tq} \sim 0{\cdot}42$, whilst for 4358 Å the latter factor is 0·59, exceeding the limit of 0·5. Thus our analysis may only be correctly applied to 5461 Å. In determining K_0 for this line we use the value of A_{st} as given by Schouten and Smit (1943) reduced by one-sixth as suggested by Elenbaas (1951, p. 37), thus $A_{st} \sim 6 \times 10^7$ sec^{-1}, and since $R = 2{\cdot}05$ cm and $\Delta\nu = 2{\cdot}52 \times 10^{10}$ sec, we find $p = 39$. From the radial temperature distribution (Elenbaas 1951, p. 73) it is possible to calculate $F(z)$ along the discharge diameter for both the emitting and the absorbing states. Since the main approximation made in the analysis is that $F(z)$ is the same for both states, use of either function in eqn. (10) should produce a negligible change in I/I_e if the approximation is valid. Calculation gives $\int_0^1 F(z)\,dz$ for the emitting and absorbing states as 0·247 and 0·276 respectively, yielding corresponding values of I/I_e from eqn. (10) of 0·306 and 0·292. The insensitivity of the absorption to $F(z)$ shown here is independent of the spectral profile. Thus similar calculations for a Doppler profile using eqn. (11) give the I/I_e values as 0·113 and 0·103, whilst a linear profile gives 0·10 and 0·09. It is clear from these calculations that the absorption is indeed insensitive to the change in $F(z)$ between the emitting and absorbing states, thus indicating the validity of the approximation made.

It is evident, however, that the magnitude of the absorption is appreciably affected by the form of the spectral profile. It is thus possible to estimate the profile at 5461 Å by comparing the theoretical values of absorption obtained for different profiles with the absorption experimentally determined by Elenbaas (1951, p. 145). In his experiment Elenbaas uses two discharges L$_1$ and L$_2$ and a monochromator arranged to measure only the radiation emitted by the discharge diameters. By measuring the intensities of radiation from L$_1$ and L$_2$ separately and together, denoted by I_1 and I_2 and $I_{1,2}$ respectively, a measure of the absorption along the diameter of L_2 of the radiation received from the diameter of L_1 is given by $\alpha = \{I_1 + I_2 - I_{1,2}\}/I_1$. Since sufficient data are available only for the Hg discharge (0·88 atm, 6 amp) we will consider only the case when L_1 and L_2 are similar discharges.

If the radiation received from L_1 is $(I_\nu\,d\nu)_1$ given by eqn. (9), then after passage through L_2 it is reduced to the value $(I_\nu d\nu)_1 \exp[-2K_0(\psi)_0{}^R]$, so that the amount absorbed is $(I_\nu\,d\nu)_1[1-\exp\{-2K_0(\psi)_0{}^R\}]$. Thus

$$\alpha = \frac{\int_{-\infty}^{+\infty}[1-\exp\{-2K_0(\psi)_0{}^R\}]^2\,d\nu}{\int_{-\infty}^{+\infty}[1-\exp\{-2K_0(\psi)_0{}^R\}]\,d\nu} = \left[2 - \frac{\int_{-\infty}^{+\infty}[1-\exp\{-4K_0(\psi)_0{}^R\}]\,d\nu}{\int_{-\infty}^{+\infty}[1-\exp\{-2K_0(\psi)_0{}^R\}]\,d\nu}\right]. \qquad \ldots\ldots(14)$$

This is the equivalent of the line absorption of Ladenburg and Reiche (1913). For a resonance profile we obtain from (14) where $a = (2/\pi)\int_0^1 F(z)\,dz$

$$\alpha = 2\left[1 - \frac{\epsilon^{-2ap}\{J_0(2iap) - iJ_1(2iap)\}}{\epsilon^{-ap}\{J_0(iap) - iJ_1(iap)\}}\right]. \qquad \ldots\ldots(15)$$

As before, $p = 39$ and with either 0·247 or 0·276 for $\int_0^1 F(z)\,dz$ we obtain from this equation $\alpha = 0{\cdot}57$. This compares favourably with the experimentally determined value of 0·54.

Closer agreement than this is not to be expected because of uncertainties in the transition probability value, the actual profile form and half-width and because of the neglect of spatial variation of profile and unknown accuracy of the intensity measurements. These uncertainties make reasonable a value of $p \sim 18$, which for a resonance profile gives full agreement between experiment and theory. The calculated values for α for other profiles are widely different from the experimental value of 0·54; thus a Doppler form gives 0·90 whilst a linear form gives 0·97. The calculations thus indicate that the line 5461 Å has a profile corresponding to a resonance form, a conclusion in agreement with that obtained from considerations of the broadening processes in the discharge.

ACKNOWLEDGMENTS

The author wishes to express his thanks to Dr. J. D. Craggs for discussions and for interest in the work. He is indebted also to Professors J. M. Meek and F. J. Teago for their encouragement. The author is deeply obliged to Dr. W. Elenbaas for permission to include in this paper the special measurements he made of the mercury half-widths.

REFERENCES

Bartels, H., 1949, *Z. Phys.*, **125**, 597, **126**, 108 ; 1950, *Ibid.*, **127**, 243, **128**, 546.

Cowan, R. D., and Dieke, G. H., 1948, *Rev. Mod. Phys.*, **20**, 418.

Edels, H., 1950, British Electrical and Allied Research Association Report L/T 230 ; 1951, *Proc. Phys. Soc.* B, **64**, 354.

Elenbaas, W., 1951, *The High Pressure Mercury Vapour Discharge* (Amsterdam : North Holland).

Fischer, E., and König, H., 1938, *Phys. Z.*, **39**, 313.

Haslam, R. N. H., 1935, *Proc. Roy. Soc.* A, **150**, 338.

Kruithof, A. M., 1943, *Thesis*, Utrecht.

Kruithof, A. M., and Smit, J. A., 1944, *Physica*, **11**, 129.

Ladenburg, R., 1930, *Z. Phys.*, **65**, 200.

Ladenburg, R., and Reiche, F., 1913, *Ann. Phys., Lpz.*, **42**, 181.

Lindholm, E., 1945, *Ark. Mat. Astr. Phys.* A, **32**, No. 17.

Mitchell, A. C. G., and Zemansky, M. W., 1934, *Resonance Radiation and Excited Atoms* (Cambridge : University Press).

Ornstein, L. S., and Brinkman, H., 1934, *Physica*, **1**, 797.

Rompe, R., and Schulz, P., 1938, *Z. Phys.*, **108**, 654.

Schouten, W., and Smit, J. A, *Physica*, **10**, 661.

Schulz, P., 1938, *Z. Tech. Phys.*, **19**, 585 ; 1948, *Ann. Phys., Lpz.*, **3**, 280.

Wendt, G., and Wetzel, R. A., 1916, *Ann. Phys., Lpz.*, **50**, 419.

Ring Shims for Coned Magnet Polecaps

By E. R. ANDREW and F. A. RUSHWORTH
Department of Natural Philosophy, The University, St. Andrews, Scotland

Communicated by J. F. Allen; MS. received 4th June 1952

ABSTRACT. The calculation of ring shims of the kind suggested by Rose has been extended to include polecaps in the form of truncated cones. Design data are presented in graphical form. Actual and predicted performances are compared for one magnet.

Very homogeneous magnetic fields are required for the investigation of certain physical phenomena. An example is the study of nuclear magnetic resonance absorption, which usually calls for a field of several kilogauss, homogeneous within at least 1 part in 10^4 over a volume of 1 cm^3. It is customary to meet these requirements by means of a magnet equipped with plane parallel soft-iron polecaps provided with ring shims of the kind calculated by Rose (1938).

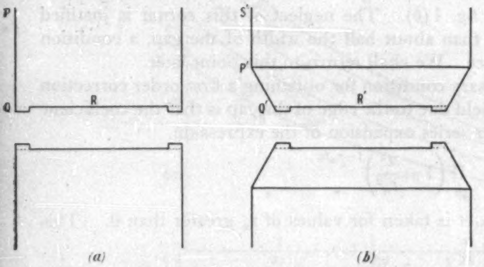


Fig. 1. Diametrical sections of typical magnet polecaps.

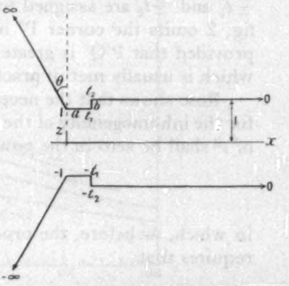


Fig. 2. Conformal mapping of the t-plane on the ζ-plane. The real t axis coincides with the surfaces of the polecaps.

Rose's analysis envisages polecaps having axial symmetry, a diametrical section of which is shown in fig. 1 (*a*). The raised rim round the edges of the polecaps compensates to a first order for the radial fall in field towards the edges of the gap. Rose gives in graphical form the height of the necessary raised rim in terms of its width. His data apply only to the case shown, namely that for which the angle PQR is 90°.

When a permanent magnet is employed, the efficient use of the permanent magnet material frequently requires the diameter of the poles to be greater than the diameter of the gap. The polecaps must then be in the form of truncated cones (fig. 1 (*b*)). The angle P′Q′R′ is now greater than 90°, so that Rose's data are not strictly applicable. The present paper gives the results of an extension of the calculation to polecaps of general angle.

It will not be necessary to repeat here Rose's analysis in so far as it is general. It will suffice to outline the method and state his results. He makes two assumptions: (*a*) that the reluctance of the soft iron of which the polecaps are made is negligible: the surfaces of the polecaps may then be treated as equipotential surfaces; (*b*) that the diameter of the polecap is large compared with the width of the gap: the potential problem then approximates to one in two dimensions, and may be solved by conformal mapping using the Schwarz transformation method (see for example Kottler 1927 or Green 1948).

Cartesian coordinates are chosen as shown in fig. 2. The x axis lies in the median plane of the gap, and the z axis is the normal to the median plane through the edges of the polecaps. Half the width of the gap is taken as the unit of length. A complex variable $\zeta = x + iz$ and an auxiliary variable t are introduced, the relationship between which is given by

$$\frac{d\zeta}{dt} = -\frac{2}{\pi t} \prod_n \left(1 - \frac{t^2}{t_n^2}\right)^{\alpha_n/\pi - 1}, \qquad \ldots\ldots(1)$$

where the t_n are the values of t (all real) at the corners of the polygon formed by the air gap, the α_n are the internal angles of the polygon, and the product is taken over the corners for which $t_n > 0$. Three values of t may be disposed arbitrarily (say 0, 1, ∞) as shown in fig. 2. The values -1 and $-\infty$ follow from the symmetry of the system about the median plane of the gap. Values of $\pm t_1$ and $\pm t_2$ are assigned to the remaining corners as shown. The diagram of fig. 2 omits the corner P′ of fig. 1 (*b*). The neglect of this corner is justified provided that P′Q′ is greater than about half the width of the gap, a condition which is usually met in practice. We shall return to this point later.

Rose shows that the necessary condition for obtaining a first-order correction for the inhomogeneity of the field due to the edge of the gap is that the coefficient of t^2 shall be zero in the power series expansion of the expression.

$$\prod_n \left(1 - \frac{t^2}{t_n^2}\right)^{1 - \alpha_n/\pi}$$

in which, as before, the product is taken for values of t_n greater than 0. This requires that

$$\sum_n \left(1 - \frac{\alpha_n}{\pi}\right) t_n^{-2} = 0. \qquad \ldots\ldots(2)$$

We will now apply this requirement to the case shown in fig. 2. The angle θ is the semi-angle of the truncated cones which form the polecaps, and will be expressed as $\eta\pi$. The width and thickness of the rim are a and b respectively. The parameters of the corners are therefore:

ζ	t_n	α_n
$i(1-b)$	1	$\pi(3/2-\eta)$
$a+i(1-b)$	t_1	$3\pi/2$
$a+i$	t_2	$\pi/2$.

Substituting these values of t_n in (2) we find

$$t_2^2 = t_1^2/[1 + (1-2\eta)t_1^2]. \qquad \ldots\ldots(3)$$

The value of a is obtained by integrating (1) from 1 to t_1. This gives

$$a = \frac{2}{\pi}\int_{t_1}^{1} t^{-1}(1-t^2)^{\frac{1}{2}-\eta}\left(\frac{t^2}{t_1^2}-1\right)^{1/2}\left(\frac{t^2}{t_2^2}-1\right)^{-1/2} dt. \qquad \ldots\ldots(4)$$

The value of b is obtained by integrating (1) from t_1 to t_2. This gives

$$b = \frac{2}{\pi}\int_{t_2}^{t_1} t^{-1}(1-t^2)^{\frac{1}{2}-\eta}\left(1-\frac{t^2}{t_1^2}\right)^{1/2}\left(\frac{t^2}{t_2^2}-1\right)^{-1/2} dt. \qquad \ldots\ldots(5)$$

For a given η and for each value of t_1 between 0 and 1, values of a and b are obtainable from (3), (4) and (5). Eliminating the auxiliary variable t, the thickness of the rim b is found in terms of its width a.

The integrals in (4) and (5) have been computed for semi-angles θ equal to 0°, 15°, 30°, 45° and 60° (η equal to 0, $\frac{1}{12}$, $\frac{1}{6}$, $\frac{1}{4}$, $\frac{1}{3}$). For each θ, values of a and b were obtained for t_1 equal to 0·4, 0·5, 0·6, 0·8 and 1·0. In fig. 3 the thickness of the rim b is plotted as a function of its width a for the five values of θ. Figures 4 and 5 show respectively curves of a against θ for constant b and of b against θ for constant a, the values being obtained from fig. 3 by interpolation.

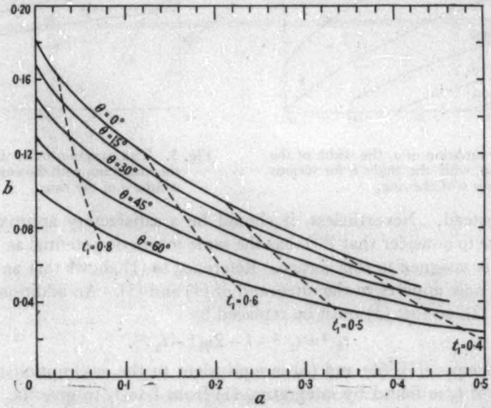


Fig. 3. Dimensions of ring shim. The width of the ring a is plotted against its thickness b for various θ, the semi-angle of the truncated cone. The unit of length is half the width of the gap. The broken lines are contours of the parameter t_1.

The integrals were evaluated using Simpson's rule. The integrand in (4) is finite throughout the range of integration. In (5) however the integrand becomes infinite at the lower limit t_2. Simpson's rule was therefore applied from t_1 down to t_0, a value slightly greater than t_2. For the remaining range of integration from t_0 to t_2 all the factors in the integrand except the last were treated as constant, t taking the value t_{m} appropriate to the mid-point of this short range. The integral could then be evaluated in this range giving the following contribution to b:

$$(2t_2/\pi t_{\mathrm{m}})(1-t_{\mathrm{m}}^2)^{\frac{1}{2}-\eta}(1-t_{\mathrm{m}}^2/t_1^2)^{1/2}\ln\left[t_0/t_2+(t_0^2/t_2^2-1)^{1/2}\right]. \qquad \ldots\ldots(6)$$

When a magnet is designed the value of b is often determined by consideration of the clearance required in the gap. If the designer wishes to calculate his own shim dimensions using (4) and (5) rather than by interpolating from fig. 3, it is necessary for him to find by trial the value of t_1 which gives the desired value of b, and then to use this value of t_1 to find the value of a. In order to facilitate the finding of the value of t_1 in such cases, lines of constant t_1 have been drawn in fig. 3 (broken lines).

We must now return to a consideration of the effect of neglecting the corner P′ of fig. 1 (b). The surface P′S′ beyond the corner may not be the surface of soft iron of low reluctance, but may for instance be permanent magnet material instead. Nevertheless, it should be a satisfactory approximation for our purpose to consider that P′S′ has the same magnetic potential as the polecap. A value t_3 is assigned to this corner. Reference to (1) shows that an extra factor $(1-t^2/t_3^2)^\eta$ now appears in the integrand of (4) and (5). An additional term also appears in (2), so that (3) must be replaced by

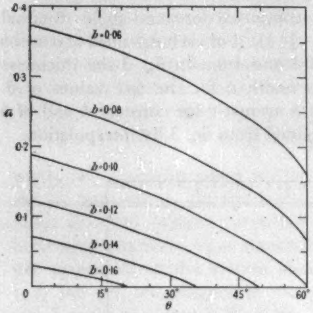


Fig. 4. The variation of a, the width of the ring shim, with the angle θ for various thicknesses b of the ring.

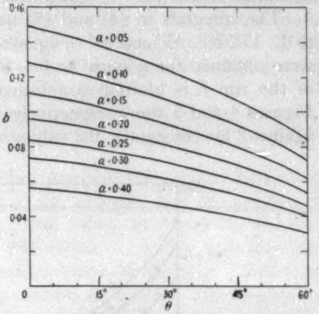


Fig. 5. The variation of b, the thickness of the ring shim, with the angle θ for various widths a of the ring.

$$t_2^{-2}=t_1^{-2}+1-2\eta(1-t_3^{-2}). \qquad \ldots\ldots(7)$$

Our use of eqns. (3), (4) and (5) is equivalent to the assumption that $t_3^{-2}\ll 1$. The value of t_3 is found by integrating (1) from 1 to t_3 to give the value of the length r of the conical surface P′Q′

$$r=\frac{2}{\pi}\int_1^{t_3} t^{-1}(t^2-1)^{\frac{1}{2}-\eta}\left(\frac{t^2}{t_1^2}-1\right)^{1/2}\left(\frac{t^2}{t_2^2}-1\right)^{-1/2}\left(1-\frac{t^2}{t_3^2}\right)^{\eta} dt. \qquad \ldots\ldots(8)$$

Approximate evaluation of (8) for various η shows that for r equal to 1·5 gap units $t_3\sim 7$, the actual value depending somewhat on η and t_1. The effect of the terms in t_3^{-2} is in fact negligible for $r\geqslant 1{\cdot}5$. For $r=1$ it is found that $t_3\sim 4$, and the terms in t_3^{-2} become appreciable. For a fixed value of a, the correction to be applied to b increases with the angle θ; at $\theta=0°$ the correction is of course zero, while at $\theta=60°$ b should be about 5% larger than the value given by fig. 3, an increase of the order of 0·003 half-gap unit. The error caused by treating the problem as two-dimensional may often be of this order.

Thus, as stated earlier, provided r exceeds about one half-gap unit the data given in figs. 3, 4 and 5 should be satisfactory. If r should be less than this, the problem must be worked out specially. The value of t_3 which gives the correct value of r for the particular θ must be found by trial from (7) and (8). The values of a and b may then be found from (4) and (5) after inclusion of the additional factor $(1-t^2/t_3^2)^\eta$ in both integrals, using the new value of t_2 from (7).

A ring shim of the type described here has been incorporated in a large permanent magnet designed by the present authors. A brief description of this magnet has been given by Hadfield and Mawson (1952), who were responsible for its construction. The soft-iron polecaps are in the form of truncated cones of semi-angle 38° 40′, having a thickness of 2·5 cm and faces of diameter 20 cm. The magnet was designed to have a gap of 5·40 cm. The thickness of the ring shim was chosen to be 2 mm, and for this thickness the appropriate width was found to be 5·43 mm. After construction the magnet was found to have a gap of 5·54 cm instead of the specified value of 5·40 cm; the dimensions of the shim should therefore have been 2·6% larger.

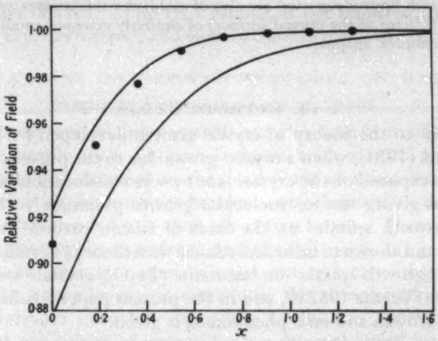


Fig. 6. Comparison of actual and predicted performances for a magnet having $\theta=38°\,40'$, and a poleface diameter to gap width ratio of 3·6. The relative variation of the magnetic field in the median plane is plotted against distance x (in half-gap units) from the edge of the gap. Circles: experimental points. Upper curve: predicted variation with shim Lower curve: predicted variation without shim.

The field in the median plane has been measured by search coil and fluxmeter and in fig. 6 is shown plotted against x, the distance inwards from the edge of the gap. The variation expected theoretically both with and without the ring shim has been calculated in the manner described by Rose and is also shown in fig. 6. While the agreement with the upper curve is not perfect, it is clear that the first order of inhomogeneity has been very largely removed. This is particularly gratifying, since it might have been doubted whether the ratio of polecap diameter to gap-width, 3·6, was sufficient to satisfy assumption (*b*) adequately; the relatively small value of this ratio may in fact be the main cause of the discrepancy between the experimental points and the upper curve in fig. 6.

REFERENCES

Green, S. L., 1948, *Theory and Use of the Complex Variable* (London: Pitman).
Hadfield, D., and Mawson, D. L., 1952, *Brit. J. Appl. Phys.*, **3**, 199.
Kottler, F., 1927, *Handbuch der Physik*, **12**, 480 (Berlin: Springer).
Rose, M. E., 1938, *Phys. Rev.*, **53**, 715.

Observations on Growth and Etch Phenomena on Haematite (Fe_2O_3) Crystals

BY A. R. VERMA
Royal Holloway College, University of London

Communicated by S. Tolansky; MS. received 2nd April 1952, *and amended 29th May* 1952

ABSTRACT. New observations of growth and etch features on haematite (Fe_2O_3) crystals are reported. Of the many growth patterns observed, the simplest are the elementary 'molecular growth spirals' in accordance with the theory of Burton, Cabrera and Frank. The step height between the successive arms of an elementary spiral has been measured by the application of multiple beam interference fringes and is shown to be equal to the x-ray value of the repeat distance. More complex growth patterns are illustrated and discussed. The etch figures reported are also of molecular thicknesses and consist of a very large number ($\sim 10^7/\mathrm{cm}^2$ of the crystal surface) of similarly oriented small triangular figures, and some more complex shapes.

---

§ 1. INTRODUCTION

ACCORDING to the theory of crystal growth developed by Burton, Cabrera and Frank (1951), when a crystal grows due to the presence of dislocations the steps exposed on the crystal face by screw dislocations wind themselves into spiral forms giving rise to 'molecular growth pyramids'. The observation of numerous 'growth spirals' on the faces of silicon carbide crystals has been reported earlier and shown to be in accordance with theory (Verma 1951). Similar observations of 'growth spirals' on haematite (Fe_2O_3) crystals were reported in a preliminary note (Verma 1952 b), and in the present paper a fuller account of the observations of growth and etch phenomena is given.

§ 2. TECHNIQUE OF OBSERVATION

Most of the observed features are of molecular height and are consequently faint. The technique found most suitable for their observation is that previously employed for silicon carbide crystals (Verma 1951) and consists of depositing a highly reflecting layer of silver on the crystal surface by thermal evaporation and then examining it in reflection using phase-contrast microscopy.

§ 3. STRUCTURE OF HAEMATITE AND ITS MODE OF GROWTH

The structure of haematite may be referred to rhombohedral axes making an angle of 85° 42′ with each other or to hexagonal axes with an axial ratio $c/a = 1\cdot36$. The structure assigned to haematite by x-ray measurements (Bragg 1924, Pauling and Hendricks 1925) may be pictured as follows. A series of rhombic cells is formed by slightly extending cubic cells along a trigonal axis so that the angle between the edges meeting in this axis is 85° 42′ instead of 90°. At every cell corner is placed a unit consisting of two Fe atoms and three O atoms. The two Fe atoms are separated by a distance of about 2·88 Å and lie at equal distances from the cell corner along the trigonal axis which passes through it. The three oxygen atoms surround each cell corner at a distance of about 1·47 Å and lie in a plane through the corner at right angles to the trigonal axis, i.e. in the (111) plane. The equilateral triangles formed by the three O atoms are oriented at 60° to each other in successive layers, i.e. are Δ and ∇ in orientation. The distance between successive oxygen layers is nearly 2·3 Å. Figure 11* shows the structure projected on the (111) plane. The structure repeats itself along the trigonal axis after *six* such (111) layers, giving the repeat distance equal to nearly 14 Å. The rhombic unit cell is outlined by joining the Fe pairs '1' to '3'; '3' to '5'; '5' to '7'; '7' coinciding in projection with '1'. This structure gives an approximate hexagonal close packing for oxygen atoms.

Haematite, which is formed in various ways and is found in rocks of all ages, is considered to occur in connection with volcanic activities as a sublimation product, often as small thin crystal plates which are parallel to the *c* axis (Dana 1948). At a temperature of about 2 000 °c, at which ferric oxide can be sublimed in crystalline form, it becomes dissociated into a magnetic oxide and consequently it will be the magnetic rather than the ferric oxide that will volatilize at these temperatures. Haematite is probably produced by the action of water vapour upon ferric chloride with the generation of HCl. The HCl vapour, if not removed, can act strongly on the haematite crystals formed, and most of the surfaces will easily etch.

§ 4. OBSERVATIONS ON GROWTH FEATURES ON HAEMATITE

Simple or Elementary Type of Spiral

On the haematite crystals many growth patterns have been observed, some of which are of the simple type. Figure 1 shows such a spiral with elementary Burgers' vector (see § 5) observed on the basal plane of a thin hexagonal plate of haematite. Originating from a right-handed screw dislocation, the spiral is seen to terminate on a left-handed one after about two turns. Thereafter successive closed loops are generated in accordance with theory.

This spiral with straight edges indicates that the rate of growth depends strongly upon the crystallographic directions, the straight edges marking the directions of minimum growth. According to Burton, Cabrera and Frank's theory (1951) the conditions of growth for a polygonal spiral are such that the 'Frenkel kinks' or exchange sites are few and far between, whereas the distance moved by the adsorbed molecule is small.

The trigonal symmetry of the spiral in fig. 1 follows from fig. 11, since for spirals on the basal planes (111) we are looking on the edges of 'monomolecular layers' along the triad axis. Further, the spiral (fig. 1) is observed to be triangular at the centre, becoming hexagonal after one or two turns. This is represented in fig. 12, which is a projection on the (111) plane, remembering that the successive arms of the spiral lie on successive 'molecular layers'.

Figure 2 is another example, of which the central growth pyramids are shown enlarged in fig. 3. Here two right-handed screw dislocations close to one another each give rise to a growth pyramid with very close spacing, indicating that the growth of these pyramids took place at quite a high supersaturation. (That these features are elevations, i.e. pyramids, has been confirmed by the application of multiple-beam fringes of equal chromatic order.) These two step lines join together after about one turn (fig. 2) to form a common spiral. Furthermore, the step lines are broad and kinked and the corners concave. This is to be attributed to the subsequent etching of the crystal giving rise to a general retreat of the step lines and rounding off the corners.

---

* All the figures in this paper are plates, printed at the end of the issue.

Interaction of Growth Fronts originating from Elementary Dislocations

A striking example of the interaction of the growth fronts from a number of dislocations emerging on the face of a haematite crystal is illustrated in fig. 4. This interaction contrasts with similar observations on silicon carbide crystals. Among the points to be specially noticed are: (i) the meeting of the three edges at an angle of 120° to each other at the centre; (ii) the way in which close pairs of dislocations give rise to peninsula-like plateaus at two places round the central part; (iii) in a total of about 20 dislocations, only two or three are of the right-handed type, the remainder being left-handed; (iv) by going round the compound growth pattern once we shall descend or move up by as many steps as the effective number of dislocations enclosed within the circuit.

Curvilinear and Bunched Spirals

On one crystal face a large number of growth hills were observed to be arrayed along the edge of the crystal. These growth hills were bunched and their spiral form could not be resolved clearly. If these growth hills have originated from screw dislocations this would be the simplest case of an assembly of dislocations, and would then form a complex transition surface (Burgers 1937, 1940). On the rest of the crystal face similar spiral features (see fig. 5) have been observed. These growth spirals appear to originate from multiple dislocations and might have been expected to take on hexagonal shapes with straight edges if left undisturbed. Further, in these spirals on the three alternate edges the steps dissociate into several components, giving an appearance analogous to that found in similar observations on silicon carbide crystals (Verma 1952 a, fig. 3). This observation, together with the curvilinear and bunched spirals of high visibility (see fig. 6), shows that, like Si–C crystals, it is also possible in haematite for spirals to originate from dislocations of multiple strength, which suggests that small portions of haematite crystals may exist as polytypes.

In a few cases rounded spirals which are heavily bunched have also been observed. Figure 6 is an example in which the spiral starts at the centre as a single step line of high visibility. Over small segments of the spiral the step lines have nearly straight edges, especially near the origin ; further from the origin the steps dissociate into several components. In the lower half of the figure over a dozen elementary faint step lines with the edges straight bunch together, leading to curved step lines of high visibility. This suggests that the high visibility rounded spiral of fig. 6 may be considered to originate from a dislocation of multiple strength.

In another case a very large number of dislocations (over 100), all of the same hand, originate in one region of the crystal. Groups of these step lines bunch together, leading again to steps with high visibility.

The circular type of spiral could be formed when the growth rate does not depend upon the crystallographic directions. Bunching of steps seems to facilitate the generation of curvilinear spirals.

§ 5. MEASUREMENT OF STEP HEIGHT

Multiple-beam interference fringes (Tolansky 1948) have been utilized for the measurement of step heights. The individual steps are rather close, and step heights are too small to be detected by the shift of the Fizeau fringes at each step.

However, as in previous work (Verma 1951), measurements were obtained by allowing a Fizeau fringe under high dispersion to pass over the peak of the pyramid, and, from the angle through which the fringe appears to bend at the peak, the average step height can be determined. Consider a pyramid formed of a series of parallel steps with straight edges, having a constant spacing d and step-height h: if a Fizeau fringe passes perpendicular to the series of steps on one side of the peak, it will appear to bend through an angle θ such that

$$h=\frac{d\sin\theta}{(1+3\cos^2\theta)^{1/2}}\frac{1}{X}\frac{\lambda}{2}$$

where X is the distance between the successive Fizeau fringes for wavelength λ. For small values of θ (which usually correspond to small values of h) as a first approximation this reduces to $h=\frac{1}{2}d\sin\theta(\lambda/2X)$, which is identical with the relation used for circular spirals.

This method was used for the measurement of the step height of the spiral in fig. 1. Since the separation between the successive steps in fig. 1 gradually decreases on going away from the centre, the Fizeau fringes passing on these steps were observed to be slightly curved. For the evaluation of step height in this case the Fizeau fringe picture and the microphotograph at the same magnification were superposed, so that the exact position of the fringe was known with certainty. For measuring θ two straight portions of the Fizeau fringe, one on either side of the centre of the spiral, were chosen and their corresponding spacing of steps d was used.

Using this method, the step height measurements for different independent trials were 12·4, 16·3, 11·2, 12·9, 17·2, 13·3, 16·0 Å.

Another measurement of average step height is possible with the fringes of equal chromatic order formed when the slit of the spectrograph was adjusted to pass over the centre of the image of the pyramid (fig. 3). This determination gave an average value of nearly 11·5 Å step height, but is subject to the error that if the slit does not pass exactly over the centre of the pyramid the value of step height will appear to be low.

These measurements, within experimental errors, show the step height to be equal to the X-ray value of the repeat unit along the trigonal axis. This spiral, therefore, is an elementary growth spiral, originating from a dislocation of unit Burgers' vector.

§ 6. MOLECULAR ETCH FIGURES

Very often on haematite crystal surfaces a variety of etch figures is observable. Figure 7 is an enlargement of one part showing a very large number of similarly oriented triangular figures covering the whole crystal surface. (For this observation the crystal surface was not silvered.) Sometimes, due to the overlapping of several smaller triangles, a much bigger triangular figure can be formed. Also some hexagonal figures, with a stepped structure and sometimes with a dot at the centre, can be observed in fig. 7. This observation of discrete triangular figures at different points is in accordance with the general observation that solution does not act uniformly over the whole face, but begins only at isolated points, proceeding more easily in certain directions, giving rise to regular figures. In the present case the triangular figure indicates three preferred directions of solution inclined to each other at 120°. The triangular shape of the etch figure conforms to the trigonal symmetry of the crystal face.

If etch begins favourably at certain isolated points in an otherwise continuous lattice surface it would seem that there are small scale lattice irregularities at these points, since solution takes place favourably at ridges, steps, etc. The small cracks of possibly near molecular dimension assumed by Smekal can explain the appearance of etch pits at such spots. According to the ideas of Frank, dislocations which provide steps for the growth of the crystal should also promote the development of etch pits. If, indeed, these etch pits have originated on the ends of dislocations, the surface density of these pits would give the density of dislocations. In fig. 7 this density is of the order of $10^7/\text{cm}^2$, which may be compared with the largest density of dislocations of the order of $10^5/\text{cm}^2$ observed on silicon carbide crystals (Verma 1952 a).

These triangular figures are the elementary etch figures with shallow depths, possibly of near molecular depth, a conclusion suggested by the fact that these features are not normally observable with usual bright field illumination, and some of the triangular features have a low visibility, even under phase-contrast illumination. However, because of the smallness of these features, it is not possible to apply interferometric techniques for their study. A hollow, which may be a different type of etch pit, is shown in fig. 8. This consists of a series of roughly circular steps some of which are seen to interlace with one another. The step height is approximately 100 Å. Fringes of equal chromatic order over this feature show it to be a pit whose total depth is about 1 000 Å.

Figure 9 is a further illustration ; here an anticlockwise triangular spiral is surrounded by the rugged and irregular broad step lines produced by solution. The kinks in these broad step lines have again triangular or probably hexagonal contours which may be due to some local growth taking place at these step edges.

§ 7. GROSS FEATURES

The elementary spiral and other molecular growth features were only rarely found on the haematite crystals studied. The majority of the growth features observed were gross, consisting of triangular pyramids (fig. 10) with broken and irregular steps of heights of a few microns with a flat triangular top with rounded corners. Three crystallographic directions inclined to each other at 120° appear as three directions of differential growth along the lines joining the three corners of the triangular feature to its centre. Similar observations have been reported previously (Maurice 1932).

§ 8. COMPARISON WITH SILICON CARBIDE CRYSTALS

Unlike the spirals on silicon carbide crystals, in haematite the point of emergence of the screw dislocation on the crystal surface, whence the spiral originates, is not marked by a hole, showing that in this case the core of the dislocation is not hollow. According to Frank (1951) a dislocation of Burgers' vector exceeding about 10 Å is only in equilibrium with an empty tube at its core, which gives rise to a crater at the point of emergence of the dislocation on the crystal surface. This conclusion will follow when the large strain energy near the dislocation is insufficient to overcome the surface energy which tends to close up the tube. However, for crystals where a complete dislocation is readily dissociated into a cluster of weak partial dislocations the core will not be hollow. Such is the case for the dislocations 47 Å high observed on paraffins by Dawson and Vand (1951). In haematite this does not seem to be a possible explanation. It is likely that the magnitudes of the surface and strain energies depending upon the elastic constants of the crystal are such that the core closes up.

Whereas in most cases of silicon carbide the spirals are more closely spaced near the centre than further out, in the spiral of fig. 1 the separation between the successive arms gradually decreases with increased distance from the centre. This is attributed, not to the gradually decreasing supersaturation, but to the limitation placed by the triangular boundary enclosing the spiral. Further, in accordance with the theory, the step lines of the polygonal spirals on silicon carbide are observed to have a curvature near the centre, whereas in fig. 1 the step lines are observed to be quite straight and the corners quite sharp at the centre. The gradual rounding off of the corners and of the step lines takes place only away from the centre.

ACKNOWLEDGMENTS

I wish to express my thanks to Professor S. Tolansky for his interest and for first suggesting that haematite might exhibit spirals. I am grateful to the authorities of the Mineralogical Department of the British Museum for their kind loan of the crystals. This work has been carried out during the tenure of a British Council Scholarship while on study leave from the University of Delhi.

REFERENCES

Bragg, W. L., 1924, *Proc. Roy. Soc.* A, **106**, 346.

Burgers, J. M., 1939, *Proc. Roy. Acad. Sci., Amst.*, **42**, 293, 378; 1940, *Proc. Phys. Soc.*, **52**, 23.

Burton, W. K., Cabrera, N., and Frank, F. C., 1951, *Phil. Trans. Roy. Soc.* A, **243**, 299.

Dana, E. S., 1948, *A Textbook of Mineralogy* (revised by W. E. Ford) (New York: Wiley; London: Chapman and Hall), p. 483.

Dawson, I. M., and Vand, V., 1951, *Proc. Roy. Soc.* A, **206**, 555.

Frank, F. C., 1951, *Acta Crystallogr.*, **4**, 497.

Maurice, Margaret, 1932, *Neuen Jahrb. f. Min.* A, **63**, 279.

Pauling, L., and Hendricks, S. B., 1925, *J. Amer. Chem. Soc.*, **47**, 781.

Tolansky, S., 1948, *Multiple Beam Interferometry of Surfaces and Films* (Oxford: University Press).

Verma, A. R., 1951, *Phil. Mag.*, **42**, 1005; 1952 a, *Ibid.*, **43**, 441; 1952 b, *Nature, Lond.*, **169**, 540.

On the Theory of Kinking

By F. C. FRANK and A. N. STROH
H. H. Wills Physical Laboratory, University of Bristol, Royal Fort, Bristol

MS. received 12*th May* 1952; *read before the Society at Bristol* 24*th March* 1952

ABSTRACT. The concept of a kink-band is introduced. It is a thin plate of sheared material in a crystal, transverse to a slip direction, bounded by opposite 'tilt walls' of dislocations. It is shown that if the angle of kinking shear γ exceeds a critical value (estimated to be $3\frac{1}{2}°$) the stress concentration at the edges of the band is sufficient for the creation of new dislocation pairs or loops, and that if the applied stress exceeds a certain value, inversely proportional to the square root of the diameter of the kink-band, the kink-band will grow by this process. The stress field of a kink-band is similar to that of a slip-band, a plate of deformation twin, or a crack under tangential stress. In suitable circumstances these may act as sources for each other. The successive generation of kink-bands, inside each other, from the same source, will produce a kink of the kind experimentally observed.

§ 1. INTRODUCTION

OROWAN (1942) has been foremost in maintaining that there is a mechanism of deformation of crystals, called by him 'kinking', which is quite distinct from the other mechanisms, slip and twinning. Speaking in terms of dislocations, the essential difference is that slip requires the generation of a succession of many dislocations all in the same slip plane, while kinking requires the generation of dislocation pairs in many parallel slip planes, regularly spaced a small number of atomic distances apart. In this paper we construct a somewhat idealized 'kink-band' in an otherwise perfect crystal subjected to a uniform applied stress, and show that the kink-band is able to act as a sufficiently effective concentrator of stress to extend itself by generating new dislocation pairs at its edges.

The kink-band we consider is a region between two approximately plane parallel 'walls' of edge dislocations, of the kind considered by Burgers (1939, 1940), Bragg (1940), van der Merwe (1950), and Read and Shockley (1949, 1950). The dislocations in these two walls are of opposite sign, so that the walls are forced in opposite directions, away from each other, by the applied stress: but the walls are of finite extent, and their edges attract each other.

§ 2. FIRST APPROXIMATION: PLANE WALLS

It will give us an instructive first approximation if we first consider an artificial constraint to be present, which keeps the walls strictly plane. The stress field of such a wall, of infinite extent, falls off with distance x from the wall as $e^{-x/D}$,

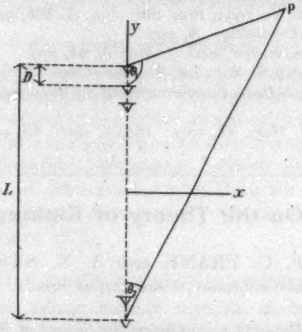


Fig. 1.

where D is the spacing between dislocations. It is negligible at distances large compared with D. A finite wall of 'height' L measured normal to the dislocation lines (infinite along their length) has a stress field like that of a large dislocation at distances large compared with L, diminishing inversely with distance; but at distances from either edge small compared with L and large compared with D the stress is independent of distance and depends only on the angular coordinate measured from the nearest edge. If (fig. 1) θ_1 and θ_2 are angular coordinates measured from either edge ($\theta_1 = 0 = \theta_2$ being the plane of the wall) the shear stress on planes parallel to the glide planes of the dislocations, at any distance greater than D from the wall, is

$$\tau_{xy} = G\lambda(\sin 2\theta_1 + \sin 2\theta_2)/4\pi(1-\sigma)D, \qquad \ldots\ldots(1)$$

where λ is the strength of the dislocations, G the rigidity modulus and σ Poisson's ratio. This formula may be obtained by approximating the sum of stresses arising from individual dislocations by an integral. In consequence of this stress, a second wall of dislocations of the opposite sign, parallel and level with the first, at a distance w from it, small compared with L, is attracted with a force

$$F = G\lambda^2 w \ln(L/w)/\pi(1-\sigma)D^2 \qquad \ldots\ldots(2)$$

per unit length measured along the dislocation lines. This result is likewise obtained by replacing the sum of the forces on individual dislocations by an integral. The maximum force is experienced by dislocations a distance approximately w from each edge, i.e. this attraction acts predominantly between the edges of the walls. Since it increases very nearly proportionately with w, the constant force $\tau\lambda L/D$ exerted on either wall by an applied shear stress τ will bring the walls into a position of equilibrium, defined by

$$(w/L)\ln(L/w) = \pi(1-\sigma)\tau D/G\lambda \qquad \ldots\ldots(3)$$

i.e. in order of magnitude,

$$w/L \sim (\tau/G)(D/\lambda). \qquad \ldots\ldots(4)$$

The resultant stress, due to the superposition of the applied stress and the two stress fields of the form (1), which tend to cancel each other except on the mid-plane between the walls, has a maximum at points on this plane from which the two neighbouring edges of the walls subtend an angle of 90°, i.e. a distance $w/\sqrt{2}$ beyond the edges of the walls. This maximum stress is

$$\tau_{\max} = \tau + G\lambda/2\pi(1-\sigma)D \qquad \ldots\ldots(5)$$

and hence, when the equilibrium condition (3) is satisfied, in order of magnitude

$$\tau_{\max} \sim \tau(1 + L/w). \qquad \ldots\ldots(6)$$

Supposing that λ/D (the angle of kinking) is of the order of magnitude 10^{-1}, $\tau_{\max}/G$ is of similar order of magnitude, and thus sufficient for the generation of new dislocation pairs; we must leave the question of whether this maximum stress occurs at the appropriate distance from the dislocations already existing till we have developed a better approximation, in § 3 below. We infer that if the angle of kinking is sufficient, the kink-band can grow by making new dislocation pairs at its edges; it can of course also shrink by mutual annihilation of pairs at the edges. To find which will happen we must discover which process lowers the energy.

The stress field of the pair of parallel tilt walls, observed at a distance large compared with D, is equal to that of two short walls bridging the gaps at the ends, as we may see most simply (fig. 2) by the consideration that the addition of two such walls of opposite sign would reduce the whole stress field to that of a slab of rotated crystal bounded by a rectangular parallelepiped of tilt walls, with a stress field effectively confined to a distance of the order of magnitude of D from them. Seen at distances further away than w, this stress field is virtually that of two strong dislocations of opposite sign, their Burgers vectors being $w\lambda/D$, directed in the plane of the main walls. The energy of the system as compared with that of the unkinked material under the same applied stress is approximately

$$U=2G(w\lambda/D)^2\ln(L/w)/4\pi(1-\sigma)$$
$$+2G\lambda^2(L/D)\ln(D/r_1)/4\pi(1-\sigma)-Lw\tau\lambda/D, \qquad \ldots\ldots(7)$$

per unit length measured along the dislocation lines; r_1 is a length, of the same order of magnitude as λ, dependent on the 'core energy' of dislocations. These three terms represent respectively the energies of macroscopic and microscopic stress fields of the dislocation array and the energy released by removing the previous state of strain from the kinked region. If, under constant τ, w/L retains its equilibrium value according to (3) this becomes

$$U=\frac{G\lambda^2L\ln(D/\lambda)}{2\pi(1-\sigma)D}-\frac{\pi(1-\sigma)\tau^2L^2}{2G\ln(L/w)}, \qquad \ldots\ldots(8)$$

which is a maximum, as a function of L, when

$$\frac{L}{\lambda}=\frac{G^2\lambda\ln(L/w)\ln(D/\lambda)}{2\pi^2(1-\sigma)^2\tau^2D}. \qquad \ldots\ldots(9)$$

Hence the kink-band will grow if it has more than a critical depth, which is inversely proportional to the square of the stress, and for $\lambda/D\sim10^{-1}$, $\lambda\sim10^{-8}$ cm and $\tau/G\sim10^{-3}$, is of the order of 10^{-3} cm.

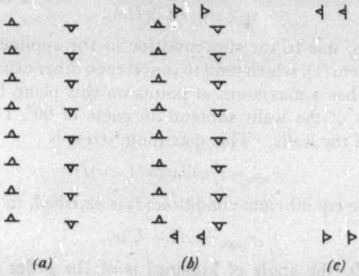


Fig. 2. (*a*) may be regarded as the superposition of (*b*) and (*c*): (*b*) has no remote stress field (representing an enclosed grain of crystal with different orientation): hence the remote stress field of (*a*) is the same as that of (*c*).

This simple treatment has brought out the main qualitative features of the situation, which we shall confirm by a more accurate but less direct method. The degree of success of the simple treatment depends on the fact that the walls have considerable stiffness, and so do remain approximately plane. Nevertheless, because the attraction between them is predominant near their edges, we must expect these edges to bend in towards each other: we shall in fact conclude that in the equilibrium configuration they bend round to become the sides of a thin elliptic cylinder.

§ 3. SECOND APPROXIMATION: THE ELLIPTICAL KINK

Consider an infinite isotropic elastic body containing a cavity in the form of an infinite cylinder, whose generators are parallel to the z-axis of coordinates and whose cross section is the ellipse

$$x^2/a^2+y^2/b^2=1, \qquad b\gg a. \qquad \ldots\ldots(10)$$

The displacements and stresses under an applied shear stress $\tau_{xy}=\tau$ have been calculated by Starr (1928). . We shall use his results, omitting terms which are of second or higher order in stress-dependence.

Suppose the cavity had been filled with a jelly of elastic moduli negligible compared with those of the external material; the applied shear stress would have produced in this jelly a uniform shear

$$\gamma=\frac{\tau(1-\sigma)}{G}\frac{b}{a}. \qquad \ldots\ldots(11)$$

Supposing the external body and the jelly both to be crystalline, with the same lattice structure in the unstrained condition, there is still atomic correspondence at the boundary of the ellipse in the strained condition. Now, suppose the jelly removed and replaced by unstrained crystalline material, identical with the external material. Since it is unstrained it produces no macroscopic stresses, and macroscopic elastic equilibrium still exists, but there is now atomic misfit at the interface. This gives rise to local centres of strain, which may be identified as dislocations; the strains due to these are only appreciable at distances from the interface of the same order of magnitude as the distance between neighbouring dislocations. The dislocation density depends on the orientation of the plug of unstrained crystal inserted into the cavity, and is a minimum when this is chosen so as to give continuity of lattice planes across the 'thickness' of the elliptical plate. The dislocations present then all have their Burgers vectors directed along the x-axis, and their distance apart in either 'wall', measured along the y-axis, is

$$D=\lambda/\gamma=Ga\lambda/\tau b(1-\sigma), \qquad \ldots\ldots(12)$$

where λ is the 'strength' of the dislocations.

Our calculation is analogous to that of Leibfried (1951) for finding the equilibrium distribution of various dislocation systems all in one slip plane, under various stress fields. Several problems of this kind were treated in terms of discrete dislocations by Eshelby, Frank and Nabarro (1951). Leibfried's treatment of the same problems by continuum theory leads to results which are asymptotically identical for the case of large numbers of dislocations, and very similar, displaying all the principal qualitative features of the equilibrium distribution, even for small numbers of dislocations. We infer by analogy that the same applies to our present problem, and that we have thus discovered, to a good approximation, an equilibrium configuration for the dislocations bounding our idealized kink-band (fig. 3). They lie on the surface of an elliptic cylinder of cross section defined by (10) and have the spacing D defined by (12). Rearranging our equations, we may say that for a constant dislocation spacing D (or constant kinking shear $\gamma=\lambda/D$) and a constant depth of kink-band $2b$ the thickness of the kink-band $2a$ varies with the applied stress τ according to the equation

$$2a=2b\tau(1-\sigma)/G\gamma \qquad \ldots\ldots(13)$$

in agreement with (4).

The radius of curvature at the edge of the kink-band is

$$\rho=a^2/b=b\tau^2(1-\sigma)^2/G^2\gamma^2. \qquad \ldots\ldots(14)$$

The stress field around the kink-band, at points not too close to individual dislocations, will be that found by Starr. In particular the shear stress τ_{xy} has a maximum at a point opposite each end of the ellipse and distant ρ from it and is there equal to

$$\tau_{\max}=(2/3\sqrt{3})b\tau/a \qquad \ldots\ldots(15)$$

which is comparable with (6).

If γ is constant, then by eqn. (11), $b\tau/a$ is constant, and hence $\tau_{\max}$ is independent of τ:

$$\tau_{\max}=(2/3\sqrt{3})G\gamma/(1-\sigma) \qquad \ldots\ldots(16)$$

(cf. eqn. (5)). If

$$\tau_{\max}\geqslant\tau_c \qquad \ldots\ldots(17)$$

the yield stress of the perfect crystal it should generate a new dislocation pair and so deepen the kink-band. A kink-band can therefore extend by the proposed mechanism if it has not less than a certain characteristic kinking shear γ_c determined by natural constants of the material:

$$\gamma\geqslant\gamma_c\equiv 3\sqrt{3}\tau_c(1-\sigma)/2G. \qquad \ldots\ldots(18)$$

If we take $\sigma=\frac{1}{3}$ we have the simple formula

$$\gamma_c=\sqrt{3}\tau_c/G. \qquad \ldots\ldots(19)$$

From Mackenzie's estimate (1949) appropriate to inert gas crystals and, for want of a better model, to metals, $\tau_c\sim G/30$, giving $\gamma_c\sim 3\frac{1}{2}°$.

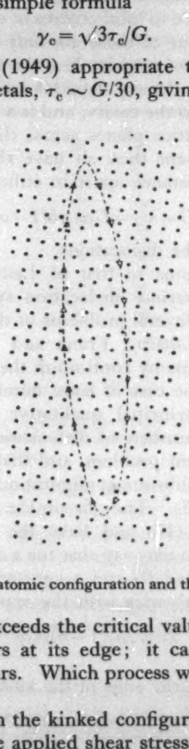


Fig. 3. Diagram shows the atomic configuration and the dislocations in a kink-band.

If the angle of kinking exceeds the critical value the kink-band can grow by making new dislocation pairs at its edge; it can in any case also shrink by annihilation of dislocation pairs. Which process will happen is decided by which lowers the energy.

The decrease of energy in the kinked configuration, relative to the unkinked configuration under the same applied shear stress, is obtained by taking Starr's (1928) value for the energy decrease (the formula for $[W]$ on p. 500 with sign reversed, neglecting the small quantity $A(\xi_0)$), and adding the 'microscopic' strain energy of dislocations in the surface. This gives, for unit length along the dislocation lines,

$$U=-\pi(1-\sigma)\tau^2b^2/2G+Gb\lambda\gamma\ln(\lambda/\gamma r_1)/\pi(1-\sigma), \qquad \ldots\ldots(20)$$

where r_1 is a length of the same order of magnitude as λ, dependent on the imperfectly known 'core energy' of a dislocation. U is a maximum, with respect to b, when

$$b/\lambda=G^2\gamma\ln(\lambda/\gamma r_1)/\tau^2\pi^2(1-\sigma)^2 \qquad \ldots\ldots(21)$$

$$\sim\tfrac{1}{2}(G/\tau)^2\gamma. \qquad \ldots\ldots(22)$$

Inserting numerical values, the smallest kink-band which will grow under a stress of $10^{-3}G$ is about 10^{-3} cm deep and 10^{-5} cm thick.

We now bring in the requirement that the kinking should continue in the same manner, with a constant kinking angle γ. The second term in (20), expressing the local energy of dislocations is represented as varying continuously with b: actually it will vary periodically, as new dislocation pairs are formed. We suppose that the configuration pauses at positions where the last pair of dislocations are a distance $D/2$ from the end of the ellipse, this being the configuration in which the strain due to the dislocation array corresponds most nearly to continuous kinking at the elliptical boundary. Then for uniform continuation we require the next pair of dislocations to be created a distance $D/2$ beyond the edge of the ellipse. If we equate this to ρ, which gives the position of maximum stress according to Starr, we have, from (14)

$$\tfrac{1}{2}D=\lambda/2\gamma=\rho=(1-\sigma)^2b\tau^2/G^2\gamma^2, \qquad \ldots\ldots(23)$$

whence

$$\tau/G=(1-\sigma)^{-1}(\gamma\lambda/2b)^{1/2} \qquad \ldots\ldots(24)$$

which is $(1-\sigma)^{-1}$, i.e. about $1\tfrac{1}{2}$, times the stress required according to (22).

We commit a logical error here by employing Starr's formula too close to the dislocation system, at a distance of the order of D instead of a distance much greater than D. However, the same conclusion is approximately confirmed by considering discrete dislocations. The position of maximum stress should not be far from the position of maximum stress from the nearest pair of dislocations, which is on the mid-plane between them at a point where they subtend an angle of 60°. If they lie on the ellipse, a distance H in the y direction from its end, their separation is $H^{1/2}b^{1/2}\tau 2^{3/2}(1-\sigma)/\gamma G$, and if H is $D/2$, the point of maximum stress due to them is a further $D/2$ beyond the end of the ellipse when

$$\tau/G=(1-\sigma)^{-1}(\gamma\lambda/3b)^{1/2}. \qquad \ldots\ldots(25)$$

We reach much the same conclusion if we require that this point of maximum stress due to the last pair of dislocations should lie on the tip of the ellipse when they are D above it: this gives

$$\tau/G=(1-\sigma)^{-1}(\gamma\lambda/6b)^{1/2}. \qquad \ldots\ldots(26)$$

Hence we conclude that there is a shear stress, dependent on the size of the kink-band as in (22), and not very much larger than is given by (22), at which the kink-band propagates with constant kinking angle.

It is necessary to enquire whether the propagation at constant angle is stable. It appears by the argument above that an insufficient stress can allow the kink-band to propagate a certain distance but with increasing kink angle, while an excessive stress could produce dislocations too far away, causing the kink angle to diminish until it is too small for propagation to continue. We may, however, use the following argument which implies that there will in general be a stable condition.

The condition $\tau \geqslant \tau_c$ defines a region S round the point of maximum stress within which the new dislocation pair will be created. In this region there will be a continual production of dislocation loops due to thermal fluctuations, but in most cases these will disappear in succeeding fluctuations. For a loop to grow under the applied stress it must be carried by the fluctuations past some critical size. Accordingly we introduce a length l, characteristic of the material and of the order of a few atomic spacings, and require that S should have linear dimensions not less than l before a new pair of dislocations may be formed.

As a first approximation we may assume that the shape of the kink-band near its end does not change as the kink-band widens under the applied stress. Then the size of the region S and its distance from the end of the band will increase in constant ratio k to one another. We have stipulated that S shall have dimensions l when the new dislocation pair is formed; it follows that this pair will be formed at a distance $l/k = h$, say, from the end of the kink-band.

We have now to consider how h will depend on the spacing D of the dislocations in the kink-band.

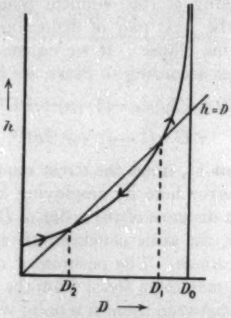


Fig. 4. Position of formation of new dislocation pair, against spacing in kink-band.

If $D > D_0 \equiv \lambda/\gamma_c$, no new production of dislocations can take place. If D is slightly less than D_0 the region S will be relatively small, that is k is small or h large. As D decreases the relative size of S steadily increases and h decreases. If h became zero for any value of D it would mean that the new dislocation pair was formed on the same glide plane as, and between, the end pair of dislocations of the kink-band. But the stress field of the latter is such as to oppose the formation of new dislocations, and since these dislocations are the most favourably situated with regard both to distance and to orientation factor we can expect them to contribute an appreciable part of the total stress field. Consequently we do not expect that h can decrease as much as to zero for any D. These considerations lead us to expect a curve of h against D of the form shown in fig. 4. The line $h = D$ either may not meet this curve, or it may intersect it in two points as shown in the figure. In the former case h is always greater than D, so that as the kink-band develops the spacing becomes larger until it eventually exceeds D_0 when no further dislocations can be produced. In the second case where $h = D$ meets the curve the points of intersection give two values of D,

D_1 and D_2 ($D_1 > D_2$) which can be continued uniformly. It is easily seen that D_2 is stable and D_1 unstable. Hence if the spacing in the kink-band is initially less than D_1 it will tend to the value D_2, while if the spacing is greater than D_1 it will increase until it exceeds D_0 when further growth of the band is prevented.

This argument is based on continuum theory. When attention is paid to the existence of a crystal lattice, we must notice that there are certain preferred angles of tilt, those for which the mean number of lattice planes between successive dislocations is an integer, which give cusped minima of surface energy with respect to angle of tilt (see Read and Shockley 1950). It is possible that stable kinking could set in on one of these preferred angles when continuum theory would not predict stability.

§4. DISCUSSION

We have dealt with a kink-band in the form of an elliptic cylinder for reasons of simplicity, in order to have a two-dimensional elastic problem, with some part of the calculation done for us in advance. In physical reality we are more likely to have kink-bands in the form of ellipsoids, roughly penny-shaped. We suppose that this case would give us more work without teaching us anything new in principle, or even making much quantitative change in the results. The remote stress field of the penny-shaped kink-band will be virtually that of a closed loop of dislocation line, roughly following its edge, but with somewhat smaller diameter, with its Burgers vector proportional to the thickness of the band.

Bounded slip-bands, kink-bands, plates of deformation, twin or martensitic transformation product, and thin cavities, cracks or notches under tangential stress, all produce stress fields of the same kind: in various circumstances any one of these may serve as the source of any other. As an example, in a substance such as rock-salt with slip planes, and slip directions at right angles to each other, a slip-band held up at any point might continue as a kink-band employing slip on the perpendicular (110) plane. The kink-bands observed by Brilliantow and Obreimow (1934) in NaCl may have such an origin. In the case of zinc a suitable initiator of a kink-band would be a small included twin of the main crystal, mechanical 'untwinning' of which would occur under a compressive stress along the base-plane.

Experimentally observed kinks are usually much thicker than the thin kink-bands of our theory. However, it is to be expected that thin kink-bands will build up into a broad kink. When a kink-band extends as far as a free surface, the attraction between the edges of the tilt walls disappears, and they should become parallel planes. A continuing stress should then force the walls apart. Now, suppose the kink-band started from some permanent source, such as a cavity or plastic inclusion. The whole force acting on the walls is available to detach them from this source and, as soon as they separate from it, it should be able to initiate a second kink-band within the first. The sequel, leading to the 'stove-pipe' kink which is typically observed, is shown in fig. 5.

The last figures in the sequence show another phenomenon: the union of two tilt walls into one. This can occur if the component dislocations are not on the same slip planes, and in general lowers the total boundary energy.

It should be noted that the slip planes on either side of a wall of a kink-band are symmetrical with respect to it. The slip planes inside and outside the band

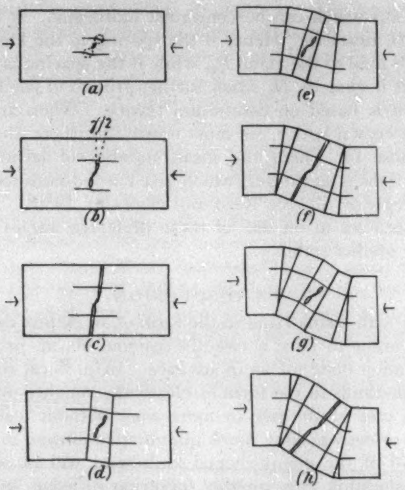


Fig. 5. (*a*) Specimen with included twin now 'untwinned' (or other stress raiser, e.g. cavity) showing stress concentration regions at its ends, and horizontal slip planes; (*b*) a kink-band has developed; (*c*) it has reached the surface; (*d*) its walls have moved apart, and a new kink band has commenced; (*e*), (*f*), (*g*), (*h*) the process continues: at (*g*) the outer walls have partly united.

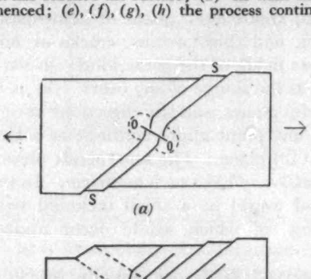


Fig. 6. (*a*) Incipient deformation band, due to slip held up at an obstacle: O, O, regions of magnified and reversed stress; (*b*) production of transverse kink-band on reversal of applied stress (assuming the initial stress does not reverse); (*c*) the growth of a deformation band.

make an angle γ with each other, so that the kink-band is tilted at an angle $\frac{1}{2}\gamma$ from the normal to the slip direction in the external material.

Figure 6(*a*) shows an incipient deformation band: a region of S-shaped lattice curvature where slip on slip planes S S is held up against some unspecified obstacle (any continuous slip through the specimen which may have by-passed the obstacle is irrelevant, and is omitted from the diagram). This accumulation of dislocations of opposite sign is somewhat similar to our kink-band, and produces regions of magnified (and *reversed*) shear stress at the points O. This may be sufficient to produce new dislocation pairs, and so the commencement of a kink-band, transverse to the slip plane, but this kink-band will not grow indefinitely because the slip in it is in the reverse direction to that directly favoured by the applied stress. It could, hypothetically, develop into a simple kink-band across the specimen on reversal of the stress, as shown in fig. 6(*b*); but under continuing stress of the same sign only a limited amount, if any, of pair creation will occur. In any case, it does not destroy the region of reversed stress. Hence dislocations on subsequent slip bands will be held up at the edge of the region of reversed stress. Thus it will propagate across the whole specimen as indicated in fig. 6(*c*) and observed by Cahn (1951). The stress field of a developing deformation band is thus very similar to that of a kink-band. The differences are that the dislocations come from outside instead of from inside, so that the kinking is in the reverse direction, and are in groups side by side on the same slip planes, so that the lattice planes are curved, instead of sharply bent, at the walls. This is a situation of higher energy, and on heating 'polygonizes' to something more nearly resembling the simple kink. Figure 6 seems to suggest that the initial obstacle must extend right through the specimen, normal to the plane of the paper, but this is unnecessary. The lattice curvature will spontaneously spread out in this direction from any small obstacle. In terms of dislocations this is a consequence of the repulsion between the similar screw dislocations which unite pairs of opposite edge dislocations on either side of the band, but its necessity is seen most simply by trying to force an S-shaped bend locally and not elsewhere in a sheet of paper.

There is more that could be said, for example with reference to twinning and martensitic transformations, and to creep, and possibly fracture, of polycrystalline metals, but we think we have said enough to indicate the considerable range of application of the theory of kinking, beyond the pure phenomenon of kinking itself.

ACKNOWLEDGMENT

We wish to thank Dr. J. D. Eshelby for helpful discussions.

REFERENCES

BRAGG, W. L., 1940, *Proc. Phys. Soc.*, **52,** 54.

BRILLIANTOW, N. A., and OBREIMOW, I. W., 1934, *Phys. Z. Sowjet.*, **6**, 587.

BURGERS, J. M., 1939, *Proc. K. Akad. wet. Amst.* **42,** 293 ; 1940, *Proc. Phys. Soc.*, **52,** 23.

CAHN, R. W., 1951, *J. Inst. Met.*, **79,** 129.

ESHELBY, J. D., FRANK, F. C., and NABARRO, F. R. N., 1951, *Phil. Mag.*, **42,** 351.

LEIBFRIED, G., 1951, *Z. Phys.*, **130,** 214.

MACKENZIE, J. K., 1949, *Thesis*, University of Bristol.

VAN DER MERWE, J. H., 1950, *Proc. Phys. Soc.* A, **63,** 616.

OROWAN, E., 1942, *Nature, Lond.*, **149,** 643.

READ, W. T., and SHOCKLEY, W., 1949, *Phys. Rev.*, **75,** 692 ; 1950, *Ibid.*, **78,** 275.

STARR, A. T., 1928, *Proc. Camb. Phil. Soc.*, **24,** 489.

LETTERS TO THE EDITOR

Supplementary Note on Ray Tracing

It seems desirable to record two small additions to the original descriptions of the methods of tracing rays through a symmetrical optical system (Smith 1920, 1945). According to the first of these methods the difference of path for a ray lying in an axial plane, which is refracted at a spherical surface, is represented by a fraction, the denominator being

$$D=4\cos\tfrac{1}{2}\psi\cos\tfrac{1}{2}\phi\cos\tfrac{1}{2}\phi'\cos\tfrac{1}{2}\psi'.$$

As the cosines of these half angles have not been found, this expression is far from convenient. The cosines of the complete angles, as well as the sines, are however already known with ample accuracy. In general the most convenient equation to use is $D=S+P/S$, where S is the sum of the four cosines and P is the product of the four sines. The second term in this expression for D is often negligible, frequently makes a contribution to the fourth decimal place only, and only occasionally amounts to a third decimal place correction to S.

In tracing rays by the second method (which is available for tracing any ray through any surface of revolution) I learn that W. Weinstein (1952) has found cases where the published procedure fails to converge. The trouble arises in finding a coordinate of the point of intersection of a ray and the surface when the equations of both are known. The equation of the surface is taken in the form $s=f(z)$ where s is the subnormal and z is the coordinate measured along the axis of symmetry. The calculation begins by guessing a value of z, say z_0. Then the point x_0, y_0, z_0 on the ray will also lie on the surface if

$$x_0^2+y_0^2=2\int s\,dz.$$

If this condition is not satisfied a new value of z is found, viz. z_1, given by

$$2s_0z_1=x_0^2+y_0^2+2\int z\,ds,$$

where s_0 is the subnormal for points in the plane $z=z_0$ and the limits of integration are sufficiently obvious. The same procedure is followed with z_1 as the assumed value of z instead of z_0, and this yields a further value of z, say z_2. In nearly all cases $z_0, z_1, z_2, \ldots$ is a rapidly converging series leading to the value of z required. But in some exceptional cases, where the ray makes a large angle with the axis of symmetry, it may diverge. The question to be considered is how the calculation should then proceed. The answer is simple : z_2 is not the best approximation that can be taken for the continuation of the computation, though in almost all cases it has been found so good that no improvement has been required. The value now suggested for further approximations will be denoted by z'. Take z_1 as a reference point for the measurement of the differences between successive approximations, and write

$$z_0=z_1+d_1 \quad\text{and}\quad z_2=z_1+d_2.$$

If the values found according to the original scheme are rapidly convergent, d_2 will be much smaller than d_1. But in any event a better approximation than z_2 is z' where

$$z'=z_1+d' \quad\text{and}\quad \frac{1}{d'}=\frac{1}{d_1}+\frac{1}{d_2}.$$

The additional labour involved in finding z' is clearly trivial, for optical workers regularly use tables of reciprocals suitable for this application. Alternatively z' is given directly in terms of z_0, z_1, and z_2 by the equation $z'=(z_0z_2-z_1^2)/(z_0-2z_1+z_2)$. It is thought that when this modified routine is adopted the point in which the ray meets the surface can always be found rapidly with whatever accuracy is required.

A somewhat extreme example illustrates this, eight decimal places being retained to show the efficiency of the modified procedure clearly.

The ray $x=0$, $0{\cdot}28y+0{\cdot}96z=1{\cdot}104$ meets the paraboloid $x^2+y^2=1{\cdot}8z$ in the point $y=1{\cdot}2$, $z=0{\cdot}8$. Suppose the approximate method is employed with the original guess $z_0=0{\cdot}81$. The next two approximations are $z_1=0{\cdot}75493878$, $z_2=1{\cdot}01925467$, showing that the original procedure leads to a strongly divergent series. These values of z_0, z_1, z_2

give $z_0'=0\cdot80050734$, and treating this as a starting value the regular routine yields $z_1'=0\cdot79768241$, $z_2'=0\cdot81062976$. Applying the new formula again to z_0', z_1', z_2' gives the improved value $z_0''=0\cdot80000089$. The series of values of z found by using the modified routine clearly converges rapidly to the correct limit.

Roselyn, Holton,
Wincanton, Somerset.
30th June 1952.

T. Smith.

Smith, T., 1920, *Proc. Phys. Soc.*, **32**, 252; 1945, *Ibid.*, **57**, 286.
Weinstein, W., 1952, *Proc. Phys. Soc.* B, **65**, 731.

Positive and Negative Joshi Effect in a.c. 'Silent' Discharges in Iodine Vapour

It is well known that the Joshi effect is associated with current pulses which can be observed when the potential difference across a suitable high resistance, placed in series with the a.c. discharge circuit, is delineated on the fluorescent screen of an oscillograph. The object of the present communication is to show some typical oscillograms obtained with an iodine vapour ozonizer tube for different applied a.c. voltages which indicate two *distinct* sets of current pulses, and to draw certain definite conclusions regarding their nature and origin. Some experimental results suggesting different origins for the two distinct sets of pulses are also described.

Figure 1 shows a few typical oscillograms obtained with an iodine vapour ozonizer, alternately exposed to and shut out from light, for a range of 50 c/s voltages. These oscillograms showed the following features with successively increasing applied voltages: (i) Starting from a low applied voltage, and critically adjusting it so as to secure just the sinusoidal 50 c/s current trace in the dark, it was found that distinct pulses appeared on irradiation, each as a line on the 50 c/s trace near each peak. The critical voltage in the particular case was 160 volts. (ii) At slightly higher voltages these pulses appeared, *even* in the dark, with an increase in number and height and persisted on irradiation. (iii) At a slightly higher critical voltage (200 volts) another set of longer pulses appeared in the dark, one, and sometimes more, near each peak of the 50 c/s current, together with the previous set of pulses. The two sets of pulses showed a distinct difference in height, position, number and distribution. On irradiation, the longer pulses were either quenched altogether or reduced considerably in height, and the shorter pulses were found to persist with an increase in number and sometimes in height. (iv) Both sets of pulses observed in the oscillograms taken in the dark increased in number and height with increasing values of the applied voltage. (v) The reduction in height of the longer pulses on irradiation was found to be gradually smaller as the applied voltage was increased. (vi) At high voltages the two sets of pulses could not be distinguished from each other, and the pulse patterns remained practically unchanged on irradiation.

We are inclined to think that the longer pulses in our oscillograms should be considered as due to discharges taking place between the negative surface charge on the glass surface of either electrode and the positive ions in a manner first suggested by Klemenc, Hintenberger and Höfer (1937), and later reported in a slightly modified form by Deb and Ghosh (1946, 1948) at the suggestion of Mitra. This view has been the basis of an explanation of the negative Joshi effect.

The other set of comparatively short pulses observed in the oscillograms should be identified with the electron avalanches formed by the Townsend collision process. With the voltage critically adjusted to obtain just the 50 c/s trace in the dark, the pulses which flash out on irradiation are considered to be due to the formation of avalanches by the photo-electrons released by the layer of negative surface charge on irradiation. With slightly higher voltages such Townsend pulses will be formed, even in the dark, by the electrons in their flight towards the anode in the normal way.* In the case of an absorbing gas or vapour like iodine it is also possible that the absorbed light energy may provide energy for accelerating the electrons sufficiently to produce Townsend avalanches.

* This view regarding the origin of the positive Joshi effect will be given in full elsewhere.

The two distinct sets of pulses—the 'discharge' pulses and the Townsend pulses—were also observed with an iodine vapour discharge tube fitted with adjustable external 'sleeve' electrodes. On examining them with a high sweep-frequency of the linear time-base of the oscillograph, each pulse of either set was found to have either a *continuous* or an *oscillatory* form of discharge according to the internal resistance value of the discharge tube. The continuous form of the Townsend discharge pulses which appeared on irradiation for a certain distance between the two electrodes when the applied voltage was near the 'threshold' value is shown in fig. 2. It can be seen that in each discharge the current at first rises, attains a peak value and then falls rapidly with time. With an increased distance between the electrodes these pulses were found to change over to the damped oscillatory form. At a certain higher voltage the 'discharge' pulses appeared in the dark as well as the Townsend pulses. By adjusting the distance between the electrodes, and by using a higher sweep-frequency, it was possible to observe the two sets of pulses, one as continuous and the other as oscillatory, or both as continuous or oscillatory discharges having distinctly different damping coefficients. The 'discharge' pulses could be easily identified as they disappeared on irradiation at some suitable applied voltage. The different origin of the two sets of pulses is strongly suggested by these observations. The duration of both sets of pulses was found to be of the order of 10^{-3} to 10^{-4} second.

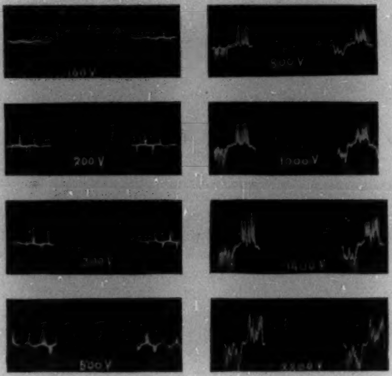


Fig. 1.

Fig. 1. Pulses in iodine vapour ozonizer at different a.c. voltages (50 c/s). The oscillograms with light-shutter closed are on the left and those with light-shutter open are on the right side.

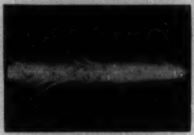


Fig. 2. Townsend pulses in iodine vapour 'sleeve'-discharge tube observed *on irradiation* at a comparatively low applied a.c. voltage with a sweep-frequency corresponding to 1500 microseconds.

It should be noted that the photo-suppression or the photo-reduction of what has been called the 'discharge' pulses is considered here as the main cause of the negative Joshi effect. It is, however, worthy of notice that the theory put forward by Joshi (1939, 1945, 1946) can explain in a general way the reduction of Townsend pulses on irradiation. The photo-dissociation theory of Harries and von Engel (1951) can also explain the same in the case of chlorine. It seems to us that these physical processes may be operative in addition to the photo-suppression of the so-called 'discharge' pulses under suitable experimental conditions.

Wireless Laboratory,
Physics Department, College of Science,
Benares Hindu University.
16th July 1952.

S. R. Khastgir.
P. S. V. Setty.

Deb, S., and Ghosh, N., 1946, *Science and Culture*, **12**, 17; 1948, *Ibid.*, **14**, 39; 1948, *J. Indian Chem. Soc.*, **25**, 449.
Harries, W. L., and von Engel, A., 1951, *Proc. Phys. Soc.* A, **64**, 916.
Joshi, S. S., 1939, *Curr. Sci.*, **8**, 548; 1945, *Proc. Indian Acad. Sci., Bangalore*, **14**, 317; 1946, *Curr. Sci.*, **15**, 281.
Klemenc, A., Hintenberger, H., and Höfer, H., 1937, *Z. Elektrochem.*, **43**, 708.

The Epitaxial Growth of Germanium on Rocksalt

A study has been made of the conditions necessary for the epitaxial growth of germanium on the (100) cleavage face of rocksalt with a view to using the method for preparing thin, single-crystal films of this material. König (1944) has examined films of germanium deposited on KBr at room temperature and finds them to be amorphous. Films which had been heated (after removal from the substrate) to 460°c–500°c were found to be polycrystalline, with large crystallites in random orientation. Substantially similar behaviour has been observed in these experiments for films deposited on rocksalt maintained at temperatures up to about 480°c.

The germanium was evaporated from a tantalum boat; to avoid the possibility of atmospheric contamination, the films were prepared in the specimen chamber of the electron diffraction camera at a pressure not exceeding 5×10^{-5} mm Hg. No trace of oxide could be detected in films so produced. Experiments were carried out at different rates of deposition of the germanium, ranging from 5 Å/sec to 100 Å/sec. The variation of beam intensity was effected by adjusting the evaporation temperature over the range 1100°c to 1300°c. Although this procedure alters the mean velocity of arrival of the atoms as well as the beam intensity, the small change in velocity over this range of temperature is unlikely to influence the deposition process. The presence of the electron beam during the condensation of the germanium was found to inhibit the deposition of the film, the thickness of the film on the part of the target traversed by the electron beam being only a few per cent of that on the neighbouring parts of the surface (fig. 1). The structure of the film is, however, unaffected by the presence of the beam.

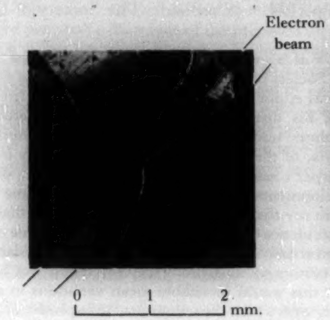


Fig. 1.

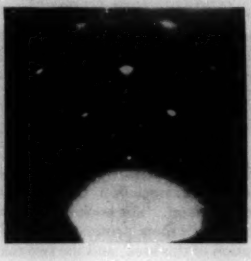


Fig. 2.

The first signs of an orienting influence of the substrate on the germanium occur at a temperature of 490°c. Sharp spots appear, superposed on continuous rings, indicating the presence of some crystalline regions with (100) and (111) faces parallel to the substrate. For (100) orientation the cube edges of the germanium are parallel to those of the rocksalt. For (111) orientation both arrangements in which the cube face diagonals coincide are found. These regions probably develop from cleavage steps on the rocksalt surface, at which the substrate influence is greater than on the plane facets. (Examination of cleaved rocksalt faces by multiple-beam Fizeau fringes shows that such surfaces are crossed by many such steps, a few tens of atoms in height and running mainly in [010] and [011] directions.) At this temperature deposition can be effected at low and high rates of evaporation. The films formed at the low rate show the reflectivity characteristic of the bulk material; those formed at the high rate are grey and lustreless and are seen to consist of large irregular aggregates. On raising the temperature to 510°c deposition is found to occur only at the high rate of evaporation, due to the appreciable evaporation of the rocksalt at this temperature. Fairly complete (100) orientation of the film occurs (fig. 2), only a trace of randomly oriented material being present. The degree of orientation is found to increase as the temperature is raised to 530°c. Above this temperature the rate of evaporation of the rocksalt is too high to allow formation of the germanium film at the highest rates of evaporation used (0·01 g/sec into a solid angle 2π; source to target distance 6 cm). The films examined ranged in thickness up to about 3000 Å; no dependence of structure on thickness was observed.

It was found to be possible to maintain a (100) rocksalt face supporting an epitaxially grown silver film at a temperature of 600°c; germanium films were deposited under this condition. It was expected that the increased surface mobility of the germanium atoms at the higher temperature would result in the formation of more perfect monocrystalline layers. This was found not to be the case (the films being similar to those grown on rocksalt at 510°c), probably due to the different nature of the substrate forces present.

University of Reading. L. E. Collins.
18th July 1952. O. S. Heavens.

König, H., 1944, *Reichsber. Phys.*, **1**, 4.

REVIEWS OF BOOKS

Glass and W. E. S. Turner, 1915–51: Articles by Twenty Contributors, edited by F. J. Gooding and E. Meigh. Pp. 144. (Sheffield: The Society of Glass Technology, 1951.) 20*s*.

Emeritus Professor W. E. S. Turner, F.R.S., celebrated his 70th birthday on 22nd September 1951. He was a founder of the Society of Glass Technology in 1916 and was its Secretary from that date until 1945. He was editor of the Journal of the Society from its foundation until 1951. In recognition of his distinguished career, twenty ex-students contributed the chapters which form the volume under review. This volume constitutes a record, even if somewhat superficial in parts, of the expansion of glass technology in this country during the period of Professor Turner's association with the industry. It should not be inferred from this that all the developments in the industry during that time have arisen from that association, but it is true to say that he had more influence on it that any other single person. The book has internal evidence of its variety of authors, and some of the contributions are more professionally written than others. This is probably unavoidable in a publication of this kind, and the two main editors obviously took the view that it was better to do the minimum of editing even though this would inevitably mean variations in style. This decision was emphasized by the urgency with which the volume had to be produced, and, as stated in the text, this is the reason why all references and diagrams have been omitted.

The range of the volume, and by implication the range of Professor Turner's activities, is considerable, covering not only glass technology but also the 'encouragement of art and design' and the 'evaluation of refractories for glass making'. The volume was obviously not intended as a textbook for experts. It is written in a style which makes it attractive to those who have a general interest in the industry or in the application of science to industry, and in no place does it become so highly technical as to be likely to cause difficulty to the more general reader. There are one or two points to which the attention of such a reader should be drawn, as the statements made may give rise to wrong interpretations. Page 38, for example, states that barium is used as carbonate or oxide; it is more usual to use the carbonate or nitrate, and the use of oxide must be extremely rare. A further statement on this page that barium flints are no longer manufactured commercially is also incorrect. On p. 58 there is a sentence which is obscure, where the statement is made that no mechanical gas-producer plant has been developed for the firing of pot furnaces. It seems to be a matter of indifference whether the gas is used for firing pot furnaces, tank furnaces, or any other type, as the point at issue is the production of a mechanical gas-producer plant, and these are well known. On p. 119, table 22, the figures for coefficient of expansion for glasses 6 and 7, namely Hysil and Monax, have apparently become transposed, the lower figure referring to No. 6 and the higher to No. 7. These however, are minor criticisms in a publication which, apart from its intrinsic interest, is a compliment to a great man, is interestingly and attractively written, and covers a very wide subject adequately in a relatively restricted number of pages.

W. M. HAMPTON.

Coloured Glasses, by W. A. Weyl. Pp. xvi + 541. (Sheffield: The Society of Glass Technology, 1951). 35*s*.

This book is the fifth monograph on glass technology published by the Society of Glass Technology under the general editorship of Professor W. E. S. Turner. The range covered by the book can best be indicated by stating that it is divided into five parts under the following general designations:

Constitution of coloured glasses.
The colours of glasses produced by various colouring ions.
The colours of glasses produced by non-metallic elements.
The colours of glasses produced by metal atoms.
Fluorescence, thermoluminescence and solarization.

In view of the range of subjects dealt with it is not to be expected that practical information immediately useful to the glass maker could be included, and it is clear that the author has written with an eye on the scientific as distinct from the production staff of the glass industry. On the other hand the approach is so specialized that it is unlikely the book will appeal to a much wider public than is provided by the glass industry, although the earlier portions, dealing with the constitution, will obviously be of value to physicists and others working in related fields. The book is very fully referenced and has two good indexes. It is well produced, with clear diagrams, but is, in the nature of things, somewhat difficult to read. The main criticism which could be levelled against it is that while the author has covered practically the whole of the literature on the subject up to the year 1949, he has done so in a rather uncritical manner. Some of the papers which are reported are of a much higher standard than others and there is no indication or comment by the author as to whether the conclusions have been modified or not by later investigators. As long as the reader will remember that he must make his own critical appreciation of the value, then this series of extended abstracts will be of considerable use. The author states that the production of the volume covered some 12 years, and the amount of information that is collected into one volume could, one supposes, scarcely have been produced in less time.

W. M. HAMPTON.

Radio Astronomy, by Bernard Lovell and J. A. Clegg. Pp. 238. (London: Chapman and Hall, 1952.) 16*s*.

The rapid growth of radio astronomy since the war has made it difficult, even for those working in related subjects, to keep in touch with the various investigations which have developed from the original work of Jansky and others in the 1930's. This difficulty is largely due to the diversity of the techniques and experience which are required in applying radio methods to astronomical investigation. Professor Lovell and Dr. Clegg have succeeded in giving a clear and authoritative account not only of the fundamental principles and possibilities of this new subject, but also of the experimental methods and results of many of the investigations which have already taken place.

The publication of this book is particularly opportune in that the first exploratory stage of the subject is perhaps now over; many preliminary studies have been made and the directions of future methods of investigation are beginning to become clear. One line of approach is illustrated by the plans, recently announced, for a large rotatable radio telescope to be built in this country under Professor Lovell's direction. The developments which have led up to these plans, and the experiments which should be possible with such an instrument, will be of interest to a wide group of readers.

No specialized knowledge either of astronomy or of radio techniques is assumed by the authors, and excellent introductory chapters are given on the fundamentals of both subjects. A detailed account is given, in the following six chapters, of the applications of pulse radar methods to the observation of meteor trails. It is shown clearly how vital has been the contribution of these methods to meteor astronomy, not only in permitting daytime measurements but in providing more accurate data than has been available from visual observations. In particular the radar measurements of meteor velocities have succeeded in settling the important question as to whether a significant fraction of meteors is of interstellar origin.

The following chapters deal with the observations and interpretation of the radio emission from the sun and the galaxy; the possibilities of investigating the structure of the solar corona by radio methods, the effects of sunspots and of solar eruptions and what is perhaps the most intriguing of the present problems, the nature of the ' radio stars '. Short accounts are given of the ' twinkling ' of radio stars due to irregularities of the ionization in the upper regions of the ionosphere, radar observations of the ionization associated with aurorae, and of the methods of investigating the moon by radar and radio methods. In a final chapter the possibility of measuring the distances of planets, and hence of determining the solar parallax, by radar methods is discussed.

Throughout the book an excellent balance has been maintained between the descriptive and the mathematical treatment. It should prove thoroughly readable for the general reader with some scientific background, yet at the same time it contains sufficient fundamental detail to satisfy those who wish to enquire more deeply into the subject. A good selection of suggestions for further reading is also given at the end of each chapter. M. R.

Electrical Phenomena at Interfaces, edited by J. A. V. Butler. Pp. vii + 309. (London : Methuen, 1951.) 32*s.* 6*d.*

The volume under review is a successor to *Electrocapillarity* by J. A. V. Butler, which appeared in 1940, and is very largely a new work in which the editor and six others have collaborated to produce a volume of very considerable interest. Electrical phenomena at interfaces 'pop up' with remarkable persistence in various fields of physics, chemistry, and biology, and it is valuable to have so many aspects of the phenomena brought together in one volume. The wide range of subject matter makes the book rather difficult to review; but careful editing and choice of topic have overcome most of the difficulties which inevitably arise from a collaborative effort of this nature.

The book opens with two chapters by J. A. V. Butler on electrical potential differences and their origin, and the electrical double-layer and electrocapillarity. There follows a discussion by J. M. Creeth of electrokinetic effects, and an account by P. A. Charlwood of electrophoresis of proteins. The electrokinetic behaviour of ionic crystals is discussed in a short chapter by T. R. Bolom, and the stability of colloidal solutions is analysed by M. B. M'Ewen, particularly in relation to the Verwey–Overbeek theories. Three chapters more in the domain of conventional physical chemistry follow : overpotential (J. O'M. Bockris), concentration polarization and the deposition of metals (J. A. V. Butler), and the behaviour of oxygen and hydrogen at electrodes (also by the editor). The last two chapters, by W. F. Floyd, bring out the importance of electrical phenomena at interfaces in biological systems. The first of these chapters discusses membrane potentials and the electrical properties of living cells, and finally there is an account of the electrical properties of nerve and muscle.

The book is well written throughout and there are many references to original papers. Altogether it should be of value to research workers and to those whose task is to teach particular features of this growing field of knowledge. J. T. RANDALL.

Polarisation de la matière. Pp. 171. 1st edn. (Paris : Centre National de la Recherche Scientifique, 1949.) 1800 fr.

This is a collection of papers on dielectric and magnetic properties of matter, read at a conference organized by the French Centre National de la Recherche Scientifique. The main emphasis is on electric polarization, and the collection forms a useful review of modern theory and experimental methods on electric dipole moments. A. H. C.

Prisms and Lens Making, by F. Twyman. Pp. viii + 629. 2nd edn. (London : Hilger and Watts, 1952.) 58*s.*

This second, enlarged edition of *Prism and Lens Making* is four times the size of the first and is a far more important contribution to optical literature. Mr. Twyman mentions in the preface that in writing it he has been assisted by contributions from other specialists, and the text contains many long extracts from original papers. These add interest and vividness without destroying the unity of the book or submerging the author's own agreeable style.

Chapter 1 is an historical introduction with many interesting quotations from original sources. It does not, however, include Leonardo da Vinci's designs for mirror-grinding and polishing machines. Chapter 2, entitled 'Single surface working', describes in helpful detail methods for the hand-making of single prisms and lenses. Chapter 3 discusses the nature of grinding and polishing and the properties of the 'polish layer'.

Chapter 4, entitled 'Tools and materials in general use', is concerned with gauges, spherometers, grinding tools, avoidance of contamination by dust during smoothing and polishing, use of wax polishers, washing and grading of abrasives, new types of abrasive, cements, plasters and varnishes used in optical work. Chapter 5, on 'Dioptric substances', discusses optical glass types, effects of thermal history on refractive index (a subject which is taken up again in Chapter 14), the working of artificial crystals and optical plastics.

Chapters 6 and 7 deal with mass production methods and with the manufacture of spectacle lenses, and Chapter 9 with that of prisms and of plane parallel plates. Chapter 8 (by R. J. Bracey) deals, rather briefly, with the making of microscope lenses.

Chapter 10 contains an excellent account of the making of non-spherical surfaces, and in particular of Schmidt corrector plates. The next chapter, on the testing of optical work, is mainly concerned with interference testing, but some account is also given of recent work on the Foucault and phase contrast tests (Zernike's name is misspelt throughout Section 227.1 and also in Section 252.3 later on). A description is given of the 'angle dekkor' and its use, as well as of some older devices. A full account of the Hilger interferometers is reserved for Chapter 12.

Chapter 13 deals with surface treatments—reflecting and non-reflecting coatings, interference filters and the making of graticules ; Chapter 14 describes the testing of optical glass and the interesting discovery that, in most practical cases, it is the disturbing effect of uneven cooling on the refractive index rather than birefringence troubles which necessitate the refinements of modern annealing.

Chapter 15 discusses the large object-glasses and mirrors needed by astronomers. The methods of H. Draper, of Sir Howard Grubb and G. W. Ritchey are described, with long quotations from the original papers, and the descriptive notes (by Mr. J. V. Thomson) on the construction of the 200-inch mirror of the Hale telescope on Mt. Palomar are accompanied by many photographic illustrations. Lastly, extracts are given from Texereau's important paper on the detection and the effects of small residual imperfections in large optical surfaces. The work concludes with a list of reference books and with four Appendices.

The volume is a goldmine of practical information solidly based on experience, while its vitality and general level of interest set a new standard for English books on this subject.

E. H. LINFOOT.

Contribution à l'étude du coefficient de viscosité des liquides, by J. HUETZ. Pp. xv + 70. (Paris : Publications Scientifiques et Techniques du Ministère de l'Air, No. 260, 1952.) 750 fr.

Contribution à l'étude aérodynamique de l'aile et de l'hélice, by M. MÉNARD. Pp. xvi + 187. (Paris : Publications Scientifiques et Techniques du Ministère de l'Air, No. 262, 1952.) 1200 fr.

Chromophotographie des vibrations d'un fluide, by J. M. BOUROT. Pp. xviii + 77. (Paris : Publications Scientifiques et Techniques du Ministère de l'Air, No. 264, 1952.) 800 fr.

Molecular Microwave Spectra Tables, by PAUL KISLUIK and CHARLES H. TOWNES, National Bureau of Standards Circular 518. Pp. vi + 127. (Washington, D.C. : U.S. Department of Commerce, 1952.) 65 c.

CONTENTS FOR SECTION A

Prof. M. H. L. Pryce. A Theoretical Attempt to Predict the Excited States of Nuclei in the Neighbourhood of ^{208}Pb 773

Mr. N. R. Steenberg. The Angular Distribution of γ-Radiation from Aligned Nuclei 791

Dr. J. A. Spiers and Dr. R. J. Blin-Stoyle. A Formulation of β-Decay Theory for Forbidden Transitions of Arbitrary Order : I—Selection Rules and Energy Spectra 801

Dr. J. A. Spiers and Dr. R. J. Blin-Stoyle. A Formulation of β-Decay Theory for Forbidden Transitions of Arbitrary Order : II—Angular Distributions . 809

Prof. C. A. Coulson and Dr. R. Taylor. Studies in Graphite and Related Compounds : I—Electronic Band Structure in Graphite 815

Dr. W. E. Duncanson and Prof. C. A. Coulson. Studies in Graphite and Related Compounds : II—Momentum Distribution in Graphite 825

Dr. R. Taylor and Prof. C. A. Coulson. Studies in Graphite and Related Compounds : III—Electronic Band Structure in Boron Nitride 834

Dr. R. McWeeny. The Diamagnetic Anisotropy of Large Aromatic Systems : IV—The Polyacenes 839

Prof. H. S. W. Massey and Dr. C. B. O. Mohr. Strong Coupling in Inelastic Collisions of Electrons with Atoms 845

Dr. H. Messel and Dr. R. B. Potts. Note on the Fluctuation Problem in Cascade Theory 854

Letters to the Editor :

- Mr. R. D. Lowde. Diffuse Reflection of Neutrons from a Single Crystal . 857
- Miss M. E. Pillow. Intensities in Band-Systems of O_2 and O_2^+ 858
- Mr. J. Sharpe. Energy per Ion Pair for Argon with Small Admixture of other Gases 859
- Mr. K. D. Bowers. Paramagnetic Resonance in Potassium Chromicyanide . 860
- Dr. D. Walker, Mr. W. T. Link and Dr. W. I. B. Smith. A Search for the Nuclear Reaction ^{4}He (α, n) ^{7}Be 861

Reviews of Books 863

Contents for Section B 865

Abstracts for Section B 866

ABSTRACTS FOR SECTION A

A Theoretical Attempt to Predict the Excited States of Nuclei in the Neighbourhood of ^{208}Pb, by M. H. L. Pryce.

ABSTRACT. ^{208}Pb has a double closed shell of 82 protons and 126 neutrons. The classification of the lowest excited states of ^{207}Pb, ^{207}Tl, ^{209}Pb, ^{209}Bi, which differ by one nucleon, should thus follow the single-particle model closely. The attempt is made to classify the lowest excited states of ^{206}Pb, ^{206}Tl, ^{206}Hg, ^{210}Pb, ^{210}Bi and ^{210}Po, which differ by two particles, in terms of the one-particle model, taking the interaction of the two particles as a perturbation. Considerable uncertainties in the order of the energy levels remain, but some definite conclusions are possible. The existence of an isomeric state of ^{210}Bi (RaE) with very high spin ($I=8$?) is predicted. This fits in with known facts about the alpha-particle emitting isomeric state.

The method can also be applied to ^{208}Tl, ^{208}Bi and the excited states of ^{208}Pb which arise by lifting one nucleon into the next shell, leaving a hole. It meets with considerable success for ^{208}Tl. Discrepancies with known facts in the case of ^{208}Pb suggest a more complex picture, according to which the first excited state would have $I=0$ and even parity. Such a state has never been observed, but its existence is not completely excluded by observation.

The Angular Distribution of γ-Radiation from Aligned Nuclei, by N. R. Steenberg.

ABSTRACT. The angular distribution of γ-radiation from nuclei aligned at low temperature is given as a series of Legendre polynomials, $I(\theta)=1+\Sigma_K A_K(T)\,P_K(\cos\theta)$ the A_K being temperature dependent. The formulation is applicable to cascade processes of arbitrary multipole orders and nuclear spins. An approximation valid for low degrees of alignment is given, as is the angular distribution from completely aligned nuclei. Application is made to ^{60}Co.

A Formulation of Beta-decay Theory for Forbidden Transitions of Arbitrary Order: I—Selection Rules and Energy Spectra, by J. A. Spiers and R. J. Blin-Stoyle.

ABSTRACT. The matrix elements of beta-decay theory are expressed in a way which not only gives the beta-spectra and angular distributions for arbitrary degree of forbiddenness, but also yields very simply the selection rules appropriate to any type of interaction.

Results for the energy spectra confirm those found by Greuling by inspection ; angular distributions and electron–neutrino correlations are considered in a subsequent paper.

A Formulation of Beta-decay Theory for Forbidden Transitions of Arbitrary Order: II—Angular Distributions, by J. A. Spiers and R. J. Blin-Stoyle.

ABSTRACT. The method of a previous paper is used to derive general formulae (valid for arbitrary Z) for the angular distributions of beta-emissions.

Studies in Graphite and Related Compounds: I—Electronic Band Structure in Graphite, by C. A. Coulson and R. Taylor.

ABSTRACT. The Bloch (tight-binding) approximation is applied to a study of graphite. The mean energy per atom and the internuclear distance are calculated in good agreement with experiment. Previous accounts are extended by including certain overlap integrals and by dealing with a set of parallel layer planes instead of only one plane. The shape of the band due to the π-electrons is calculated, and the experimental x-ray emission spectrum of graphite is interpreted. It is concluded that the role of σ-electrons is greater than is often supposed.

Studies in Graphite and Related Compounds: II—Momentum Distribution in Graphite, by W. E. Duncanson and C. A. Coulson.

ABSTRACT. The space wave functions for the π-electrons in a graphite layer, as found in Part I, are transformed into momentum coordinates, and the radial distribution function, including σ-electrons and π-electrons is calculated. Using the theory of Jauncey and DuMond the shape of the Compton profile in X-ray scattering is predicted. Both this, and the momentum distribution, are in excellent agreement with experimentally determined values. The introduction of a plausible scale factor in the atomic wave functions would make the agreement almost perfect.

Studies in Graphite and Related Compounds: III—Electronic Band Structure in Boron Nitride, by R. Taylor and C. A. Coulson.

ABSTRACT. An analysis is made of the distribution of π-electrons in the crystal of boron nitride. The insulating character of this substance is shown by the fact that these electrons completely occupy one band, of width approximately 2·4 ev, separated from the next higher band by a gap of about 4·6 ev. The density of states in the band is calculated, accurately for a single layer plane, and approximately for the three-dimensional crystal. Until the energies of the σ-electrons are better known, further progress in understanding this substance will be difficult.

The Diamagnetic Anisotropy of Large Aromatic Systems: IV—The Polyacenes, by R. McWeeny.

ABSTRACT. The π-electron diamagnetic susceptibilities of polyacenes of up to ten rings are calculated by a method introduced in an earlier paper: the method, which involves only a single numerical integration, is simplified. An alternative method of calculation, which has certain attractions and is widely applicable, is sketched in an Appendix.

The relationship between the diamagnetic anisotropy and the number of rings in the molecule is found to be almost linear.

Strong Coupling in Inelastic Collisions of Electrons with Atoms, by H. S. W. Massey and C. B. O. Mohr.

ABSTRACT. The various approximations which have been used to calculate the probability of excitation of an atom by impact with an electron whose energy is not greatly in excess of the threshold are known to fail under certain conditions. This is especially true in the important cases in which the excitation proceeds mainly through electron exchange. In these cases the exchange coupling is too strong for approximate methods, which assume weak coupling, to be valid. By using a schematic model exact solutions for strong exchange coupling are obtained and compared with the results of different weak coupling approximations. It is found that the Born–Oppenheimer approximation is never reliable at low electron energies but if distortion of the incident and scattered waves by the atomic field is allowed for correctly much better results may be obtained.

Note on the Fluctuation Problem in Cascade Theory, by H. Messel and R. B. Potts.

ABSTRACT. A general relation between two fundamental functions appearing in fluctuation problems in cascade theory is derived. This leads to a direct method of solving for the rth moments of the distribution function. It is also shown that the Jánossy *G*-equations are superfluous for the solution of the fluctuation problem.

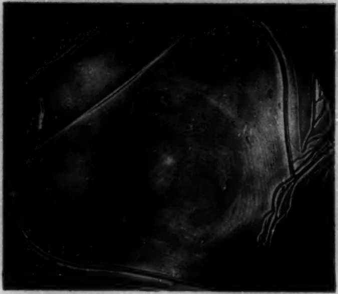


Fig. 1. ×45

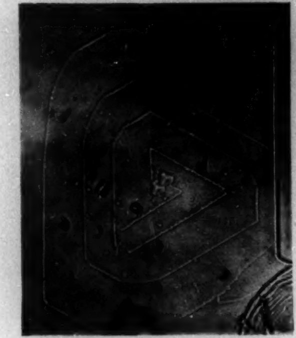


Fig. 2. ×120

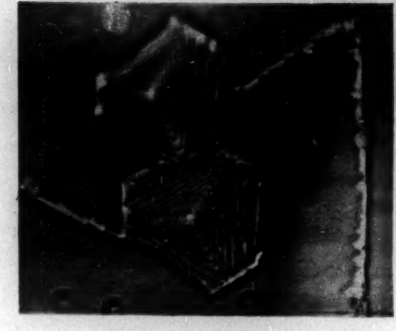


Fig. 3. ×1400

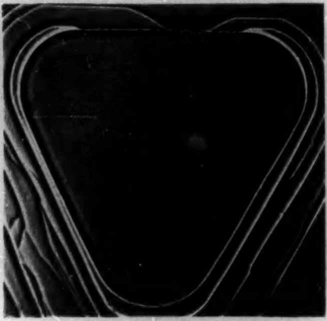


Fig. 4. ×135

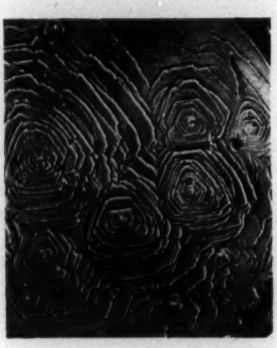


Fig. 5. ×60

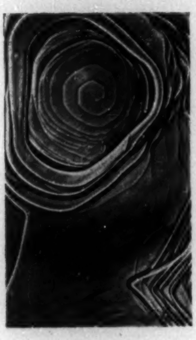


Fig. 6. ×90

PLATE I.

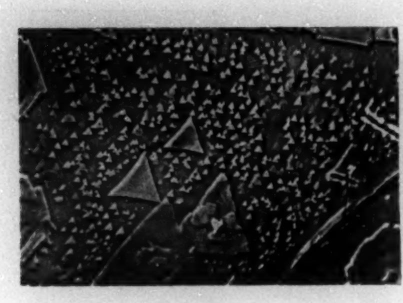


Fig. 7. ×400

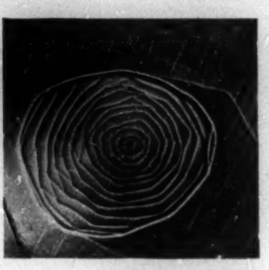


Fig. 8. ×300

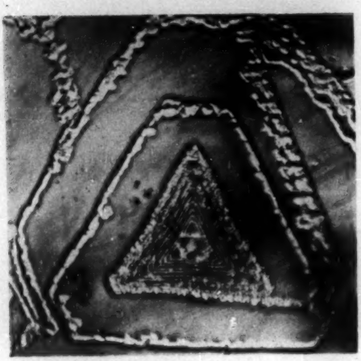


Fig. 9. ×1000

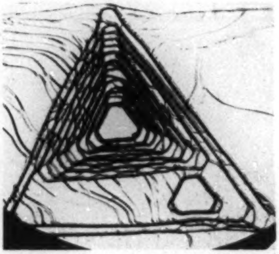


Fig. 10. ×60

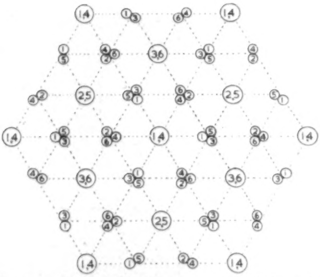


Fig. 11. Structure of haematite (Fe_2O_3) projected on (111) plane. Bigger circles represent Fe atoms and smaller circles oxygen atoms. The numbers inside the circles refer to the plane in which the atoms lie.

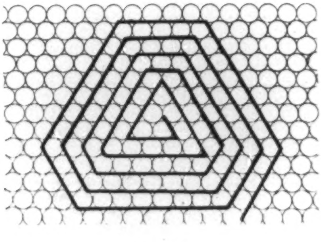


Fig. 12.

PLATE II.